Friedrich Schiller University Jena
Faculty of Chemistry and Earth Sciences
Institute of Geography
Department of Earth Observation

WS 2015/2016

Monitoring urban growth by means of multi-temporal time-series of optical Landsat data

MASTER'S THESIS
submitted in partial fulfillment of the requirements for the degree of
Master of Science in Geoinformatics

Supervisors:
Dr. Sören Hese (FSU Jena)
Dr. Mattia Marconcini (DLR-DFD)

submitted by:
Soner Üreyen
Programme: Geoinformatics MSc
Semester: 5
Matr. number: 112064
E-Mail: soner.uereyen@gmail.com

Date of submission: 24.03.2016

This master's thesis emerged at the German Aerospace Center (DLR)
German Remote Sensing Data Center (DFD)
Department of Land Surface (LAX)
Team Urban Areas and Land Management (SL)

Deutsches Zentrum
für Luft- und Raumfahrt
German Aerospace Center

Abstract

Rapid urban expansion in cities across the globe is forcing the development of appropriate methods to monitor the current status as well as the historical evolution of urban areas to support applications like urban planning and management. Here, remote sensing data have proven to be an effective tool to delineate urban areas, since they provide satellite imagery for large areas at a relatively high temporal frequency with the additional advantage of free access to data archives comprising data back to 1972. This study proposes a novel fully automated classification system based on Support Vector Machines (SVM) to derive urban extent maps. For this purpose, multi-temporal Landsat time-series namely temporal statistics consisting of selected spectral and temporal indices, have been employed for the 8 sites included in the urban supersites initiative of the Group on Earth Observation (GEO) and the investigation periods 2002–2003 and 2013–2015. The proposed methodology includes (1.) the pre-processing of Landsat scenes and calculation of the temporal statistics, (2.) the enhancement of the Global Urban Footprint (GUF), which is applied for an automated and random collection of training samples, (3.) the collection of training points for a set of configurations to overcome impacts of randomness, (4.) the application of a majority voting strategy to obtain a final urban extent map, and (5.) the implementation of an extensive accuracy assessment. The derived results report the automated SVM based classification system to be quite promising and in addition it proved to be very robust since it resulted in high accuracies throughout all study areas. In general, the obtained overall accuracy and Kappa coefficient is always higher than 91.38 % and 0.827, respectively.

Table of Contents

List of Figures

List of Tables

Abbreviations

AOI Area of Interest
ASTER Advanced Spaceborne Thermal Emission and Reflection Radiometer
ATKIS Official Topographic-Cartographic Information System
CV Coefficient of Variation
DEM Digital Elevation Model
DLR German Aerospace Center
DT Decision Tree
ETM Enhanced Thematic Mapper
GEO Group on Earth Observation
GDAL Geospatial Data Abstraction Library
GHSL Global Human Settlement Layer
GUF Global Urban Footprint
ID Index of Dispersion
JRC Joint Research Center
LiDAR Light Detection and Ranging
LST Land Surface Temperature
NDBI Normalized Difference Built-Up Index
NDGB Normalized Difference Green Blue
NDMIR Normalized Difference Middle Infrared
NDRB Normalized Difference Red Blue
NDVI Normalized Difference Vegetation Index
MNDWI Modified Normalized Difference Water Index
MSS Multispectral Scanner
OECD Organization for Economic Cooperation and Development
OLI Operational Land Imager
PA Producer's Accuracy
PALSAR Phased Array L-band Synthetic Aperture Radar
PRD Pearl River Delta
RBF Radial Basis Function
SAR Synthetic Aperture Radar
SB-04 Task 4 on Social Benefits
SFFS Sequential Forward Floating Selection
SRTM Shuttle Radar Topography Mission
SVM Support Vector Machines
TM Thematic Mapper
UA User's Accuracy
USGS United States Geological Survey

UST Urban Structure Types
VHR Very High Resolution
WRS-2 Worldwide Reference System 2

1 Introduction

Along with the drastic rise of world population, the increase of urban population is as high as never before. Indeed, it has been estimated that in 2007 the number of urban dwellers surpassed the rural population for the first time in history. In 2014 54 % of 7.2 billion humans resided in urban areas, while just 65 years earlier in 1950 only 30 % of the global population was urban. The degree of urbanization mostly depends on the region: the urban population of developed countries (including Europe and North America) is accounted for more than 73 % of the total population, whereas this share is lower than 50 % in developing countries. However, the latter are those which are nowadays experiencing the fastest growth with peaks in Asia and Africa and this trend is expected to further continue even at high speed in the next years. This is also reflected in the number of megacities which is estimated to rise from 28 in 2014 to 41 in 2030 and most of them will be in today's developing countries. In the light of this constant increase of urban population it is expected to reach a share of 66 % by 2050 of urban dwellers in comparison to the world population (UNITED NATIONS 2015:7ff.).

It is worth noting that people are moving more and more to the cities since they are centers of commercial and service industry, transportation, and government making them a driving force for economic growth. The economic attractiveness of cities is a key pull factor for the rural population seeking for employment and a life with higher quality (WAKODE et al. 2014:109). In addition, the economic importance of cities is also reflected by the share of the global gross domestic product generated in urban areas, which is estimated to be around 80 %, with increasing tendency (LALL et al. 2013:1).

In such framework, it is of great importance to evolve strategies for sustainable development of urban areas in order to meet the demands for an increasing urban population. If urbanization occurs uncontrolled it most likely leads to the absence of necessary urban infrastructure which endangers a sustainable development. Rapid urbanization has also an impact on the environment, resulting in relevant changes of the hydrological system and the climate. Furthermore, due to the higher population density compared to rural areas, cities are more vulnerable to the effects of climate change as well as natural hazards.

To address these issues reliable data capturing the current extent of urban areas as well as the human induced land transformations over time are essential. Nevertheless, due to the rapid urban growth and the complexity of urban areas it is challenging to obtain data with high accuracy and reasonable temporal and spatial resolution. In this context, remote sensing data have gained an important role, as they can cover large areas at medium (30-100 m) and even at high spatial resolution (<10 m) at relatively high temporal frequency (SCHNEIDER & WOODCOCK 2008:660).

Despite high data availability there is currently no urban map on a global scale with an acceptable accuracy and spatial resolution. In this study, an approach to detect urban areas

and monitor its spatiotemporal dynamics at a resolution of 30 m by means of Landsat data which are used together with the Global Urban Footprint (GUF) layer (ESCH et al. 2013) is proposed. The Landsat archive enables free access to optical satellite images since 1972 and is therefore an important data source for the creation of dense time-series of images, while the additionally used GUF is a worldwide map of human settlements at a resolution of 12 m. Specifically, the developed method is tested for selected study areas, which are the so-called urban supersites defined by the Task 4 on Social Benefits (SB-04) of the Group on Earth Observation (GEO) (WENG 2014a:6). The urban supersites include Athens (Greece), Atlanta (USA), Beijing (China), Istanbul (Turkey), Mexico City (Mexico), Los Angeles (USA), the Pearl River Delta (China), and Sao Paulo (Brazil).
The methodology comprises image processing, classification, and analysis of the results for the time period between 2002 and 2015. First, Landsat images are pre-processed and then multi-temporal time-series are calculated, hereafter referred to as temporal statistics. For the classification of urban areas the Support Vector Machines (SVM) classifier is employed. To verify the accuracy of the resulting maps a comprehensive validation is conducted.
The main research objectives of this study are to develop a transferable and robust classification method employing the Landsat based temporal statistics to automatically derive binary urban extent maps, to enhance the GUF by means of the temporal statistics and use it afterwards as a mask to collect training samples unsupervised, and to investigate the urban growth for the study areas between 2002 and 2015. Corresponding to these objectives the following research questions shall be investigated: (1) What are the challenges and benefits of the Landsat temporal statistics?; (2) are the Landsat based temporal statistics suitable for the improvement of the GUF?; and (3) is the developed classification approach transferable on a global scale?
This study is structured in 7 chapters. Chapter 2 introduces to the terminology and addresses the necessity for urban monitoring by means of remote sensing tools, moreover, it describes relevant principles of remote sensing, applied software tools, and gives an insight into the state of the art in urban remote sensing, focusing on application fields and recent spatial data sets on urban areas. In Chapter 3 the study areas are presented. The used data basis including the Landsat data archive, elevation data, and the GUF are described in Chapter 4. Additionally, this chapter also deals with the development of the methodology, comprising the image processing, the enhancement of the GUF, and the classification of urban area extent. The resulting urban extent maps and their accuracy are presented in Chapter 5. In Chapter 6 the obtained results are discussed, with emphasis on the suitability of the temporal statistics, the transferability of the classification approach, and the analysis of the detected urban growth. Chapter 7 concludes this study summarizing the key findings and giving a brief outlook on future developments.

2 Background

In this chapter, the focus is on relevant theoretical background and the state of the art in urban remote sensing. First, the term 'urban area' is defined and then the need for the monitoring of urban areas is explained. Afterwards, principles of optical remote sensing regarding urban applications are described followed by an explanation of classification concepts with an emphasis on Decision Trees (DT) and SVM. In addition, the functionality of the Geospatial Data Abstraction Library (GDAL) software is outlined. The state of the art in urban remote sensing is illustrated in the last section of this chapter by presenting research conducted on urban feature extraction, urban environmental analyses, and urban growth monitoring. Moreover, an overview of recent global urban maps based on remote sensing data is presented with detailed explanation of the GUF.

2.1 Definition of urban areas

Presently, there is no international agreement about what can be considered as urban area. Generally, governments use approaches similar to that defined by WEEKS (2010:34), who describes 'urban' as a function of population, spatial size, density, as well as economic and social organization. Nevertheless, there exist attempts from different organizations, like the European Commission or the members of the Organization for Economic Cooperation and Development (OECD), to harmonize the definition of urban areas. Specifically, the suggested approach is based on a spatial concept of urban centers and distinguishes between densely, intermediate, and thinly populated areas. The first class corresponds to urban centers with a minimum population of 50,000 and a density of at least 1,500 inhabitants per km²,whereas the second includes urban clusters with a minimum population of 5,000 and a density of at least 300 inhabitants per km². Instead, if a grid cell of 1 km² has less than 500 inhabitants, then it is classified as rural (DIJKSTRA & POELMAN 2014:5f.).
In this thesis, as often done in the context of urban remote sensing, the term 'urban' refers to the totality of sealed surfaces, hence including residential, industrial, and commercial zones as well as transportation infrastructure such as roads. Instead, construction and extraction sites as well as green areas within cities are categorized as non-urban.

2.2 Need for urban monitoring

Urbanized areas are estimated to cover only about ~1 % of the Earth's surface, but are home to more than half of the world's population and represent the focal point of the economic activity (SCHNEIDER et al. 2010:1733). At the same time, cities are cause of many global environmental problems, such as waste stream and pollution of air, water, and land surface (WENG et al. 2014b:15), and they represent a key source of anthropogenic carbon emissions (SVIREJEVA-HOPKINS et al. 2004:296).

In urban areas, the exposure to natural hazards and human induced risks is hence particularly high. In such context, although the precise effects of global climate change are not yet clear, it is fundamental to already define specific mitigation and adaptation strategies (CARTER et al. 2015:3). As an example, it is estimated that the urban heat island effect will further increase in the coming years with great impacts on the health and well-being of urban inhabitants. This will result in a more intensive use of air conditioners, which will in turn cause higher greenhouse emissions (GRIMMOND 2007:86f.). Other impacts of the ongoing global environmental change also include "heat waves, drought, heavy precipitation, cyclones and extreme high sea level events" (MARCONCINI et al. 2015:1); here, as the frequency of these events increases "the pressing nature of adaptation in cities becomes apparent" (CARTER et al. 2015:4). Nonetheless, despite the increasing number of researches about urban climate change and the higher awareness towards its threats, related adaptation procedures are not yet the norm. The provision of spatial data on urban areas is then central for supporting such activities, as well as other applications like urban planning and management. In this framework, mapping the extent of urban areas over time is essential for understanding their evolution and properly modeling future trends and, nowadays, it can be effectively addressed by means of Earth observation technology (SCHNEIDER 2012:689). Accordingly, the development of a robust approach capable of reliably outlining the spatial growth of human settlements is of great importance for a variety of studies.

2.3 Remote sensing

Remote sensing images collected by airborne or spaceborne sensors provide extensive information of the Earth's surface. Two main families of remote sensing sensors exist, namely active and passive. While active systems emit microwaves and register the backscattered portion of the emission, passive sensors capture reflected or emitted electromagnetic radiation from the Earth's surface. In detail, passive optical remote sensing sensors detect the visible and infrared portion of the electromagnetic spectrum; the corresponding radiance measured by the sensor depends both on the properties of the surface materials and on the specific atmospheric conditions. Examples of satellites mounted on board optical remote sensing instruments are the Landsat suite, IKONOS, SPOT, Quickbird, and Sentinel-2. Among these, the Landsat data archive is the only one providing access to images acquired with 4 to 11 spectral bands back to the year 1972. Initiated with the Landsat Multispectral Scanner (MSS), which recorded at a spatial resolution of 60 m, the Landsat mission was continued with the Thematic Mapper (TM) sensor mounted on board of Landsat 4 and 5 satellites and the Enhanced Thematic Mapper (ETM) mounted on board of Landsat 7, which record images at a resolution of 30 m and were launched in 1984 and 1999, respectively. However, the latter sensor induced discontinuities in the data archive due to the loss of the Scan Line Corrector (SLC) (WULDER et al. 2008:956f.). The data continuity was ensured with the launch of the Landsat Operational Land Imager (OLI), also referred

to as Landsat 8, in 2013. As of 1st January 2015 the Landsat archive comprised 5,532,454 images, most of them acquired by the TM, but only two years after the launch of the OLI, it already had the highest annual image acquisition rate (WULDER et al. 2015:2f.). Since the data archive is operated on a free data policy a large number of images is available for various fields of application such as land cover change monitoring and also specifically for urban change monitoring, because of both the spatial resolution as well as the relatively high temporal (the revisit time is 16 days) and geographic coverage of the data.
The already mentioned spatial resolution is one of the key characteristics of remote sensing sensors, further include the spectral and temporal resolution which are briefly explained in the following with respect to urban remote sensing applications. Here, the use of optical images at coarse resolution (> 100 m) is limited for urban monitoring because of the number of objects occurring in one pixel. The finer the spatial resolution the more objects can be separated and less mixed pixels occur. Due to highly heterogeneous environments in urban areas and objects on a spatial scale of only a few meters data with medium (10 to 100 m) and high as well as very high spatial resolution (< 10 m) are most applicable for urban monitoring (WENG et al. 2014c:51ff.).
Furthermore, a remote sensing sensor is only recording images with a selected number of spectral bands. The higher the number of spectral bands, the higher the spectral resolution. Landsat 8 records images in 9 spectral bands, within the visible and infrared spectrum. In comparison, hyperspectral sensors are able to deliver spectral signatures covering the complete spectrum with more than 400 spectral bands. The usage of hyperspectral data in urban applications would simplify the separation of different surface materials. Hence, both high spatial and spectral resolution would lead to clearly distinguishable objects and spectra (HEROLD et al. 2003a:1907). However, this configuration is not applicable on larger scales due to extremely high costs.
The temporal resolution of a sensor corresponds to its frequency of revisiting a specific area. Here, sensors with high temporal resolution are of great importance for time-series analysis which are valuable for the assessment of the “spatial and temporal dynamics of the processes of urban growth and land use change" (HEROLD et al. 2003b:287).

2.4 Image processing

2.4.1 Geospatial Data Abstraction Library

GDAL is an open source geospatial toolbox for raster and vector data processing. The GDAL project had its origin in 1998 and at that time the main objective was to establish a software for “efficient and flexible file format translation” (WARMERDAM 2008:90). Meanwhile, GDAL has become the most commonly used open source geospatial package and also the foundation for many software packages, such as QGIS, GRASS GIS, and R Statistics. The GDAL command line utilities are written in C\C++ and Python language and are available for major operating systems. The strength of this software is the provision

of a comprehensive toolbox for raster data processing. An important tool is gdal_translate, which is used to change data formats, spatial extent or resolution of raster data. Additional powerful tools are gdalwarp and gdal_merge.py. These are used to reproject raster data, mosaic images, and stack single raster bands to one raster file. The utility gdaldem enables the analysis and processing of Digital Elevation Models (DEMs) in raster format by calculating e.g. the aspect, slope, and roughness (MCINERNEY & KEMPENEERS 2015:3ff.).
There are also third party extensions to the utilities of GDAL available addressing the processing of remotely sensed image data. One example is the pktools toolbox, which is complementary to utilities of GDAL and uses a very similar interface; however, some tools are also overlapping with existing GDAL utilities. Pktools enables the collection of training samples, the classification of images, the extraction of statistics, and the implementation of accuracy assessment (MCINERNEY & KEMPENEERS 2015:173ff.). Furthermore, using pkfilter texture features can be extracted which allow to account for the relation between neighboring image pixels.

2.4.2 Classification

The classification of remote sensing images is a process of value adding to the acquired raw data and delivers information on the land cover and land use of the Earth's surface, which is essential in many applications. In the literature, several dedicated algorithms have been presented, such as the maximum likelihood classifier, artificial neural networks, DTs, and SVM (CIVCO 1993, FRIEDL & BRODLEY 1997, HASTIE et al. 2009). The latter two classifiers are further described in the following.

2.4.2.1 Decision Trees

The DT classifier is employed in remote sensing applications to assign a class label to a pixel. This method belongs to the group of non-parametric and supervised classifiers (FRIEDL & BRODLEY 1997:400). If no training data (i.e., sample data for which the corresponding information class is known a priori) is used during the classification process, then DTs are unsupervised classifiers. At this, the classification model consists of an array of tests, which are sequentially applied to the input data, starting from a root node. Each node or branch is splitting the input data to subsets until the terminal nodes (also referred to as 'leaves') are reached (PAL & MATHER 2003:556). The class labels are assigned at the terminal nodes and to find the best split at each node some splitting measures are used, which are explained more detailed in KOTSIANTIS (2013:263ff.). The definition of the logical ruleset is determined by the user or by means of training data. An advantage of DT is their clear structure, which makes the results easy to interpret.

2.4.2.2 Support Vector Machines

SVM are supervised, non-parametric and binary classifiers. In particular, they aim at separating the two given classes by means of a separating hyperplane maximizing the

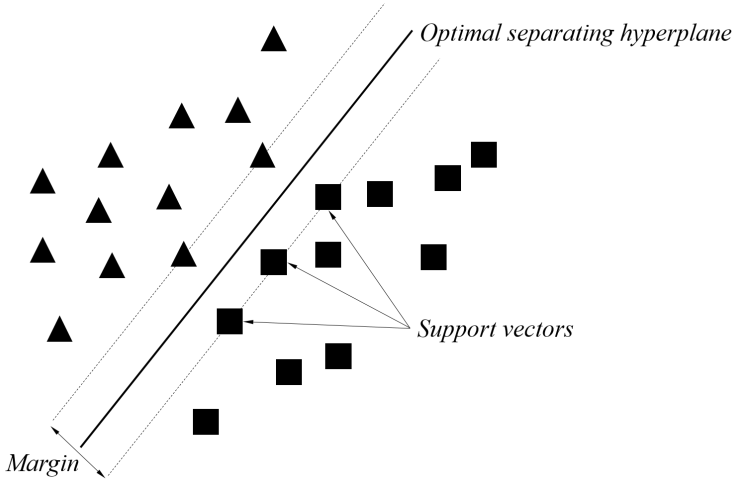

Figure 2.1: Example of linear SVM (adapted from: BURGES 1998:129).

distance between their two closest samples (i.e., margin) (HASTIE et al. 2009:132). It is possible to demonstrate that the only samples affecting the position of the hyperplane are those lying on the margin, which are called support vectors (MOUNTRAKIS et al. 2011:248).

Figure 2.1 shows an example of linearly separable classes, which can be perfectly separated by a hyperplane in a two dimensional space. However, if the classes are not linearly separable, SVM allow to map them through special 'kernel functions' into a higher dimensional feature space where they can be separated by means of a hyperplane (BURGES 1998:138). This is illustrated in Figure 2.2. Commonly used kernel functions in SVM are linear, polynomial, radial basis function (RBF) and sigmoid kernels. The kernels which are mostly applied in SVM classification of remote sensing data are the RBF and polynomial kernel (MOUNTRAKIS et al. 2011:248). The definition of the kernel width, which is one of the SVM key parameters, may lead to over fitting or over smoothing (MOUNTRAKIS et al.

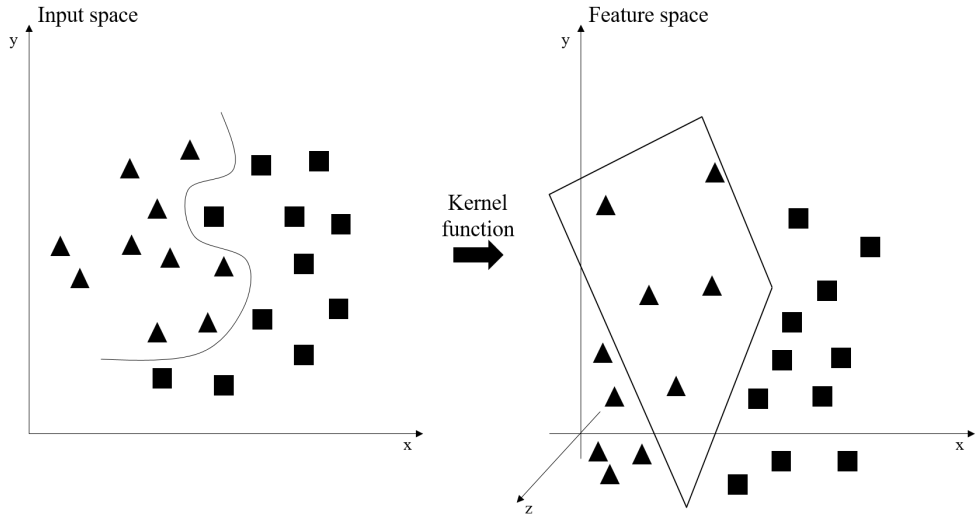

Figure 2.2: Mapping input space to a higher dimensional feature space using kernel functions (adapted from: HUANG et al. 2010:88).

2011:256). In the case of RBF kernel the parameter is referred to as gamma (γ).
Since, in real-life problems it is hard to define an optimal hyperplane which perfectly separates two classes even in the transformed high-dimensional space, SVM are allowing misclassifications in the learning phase. However, they implement a penalization cost C which is associated with the wrongly classified samples (HUANG et al. 2002:728f.; HASTIE et al. 2009:419; MOUNTRAKIS et al. 2011:256).

2.5 State of the art in urban remote sensing

In this section an overview is given on recent studies in the framework of urban remote sensing. In particular, due to the fast growing urban population and the need for data on urban areas, related research has rapidly emerged over the past decade. This is confirmed by the number of corresponding scientific publications, which increased sevenfold in the period from 2000 to 2012 (WENG 2014a:2). The following literature review will focus on methods for urban feature extraction, urban environmental research, urban growth mapping, and global urban extent maps.

2.5.1 Delineation of urban features

Urban areas exhibit a wide range of spatial and spectral characteristics, such as diverse roof types, roads, vegetation, and bare soil. Hence, an accurate mapping of urban land cover and land use requires very high resolution (VHR) remote sensing data (HEROLD et al. 2003a:1907). In case of urban spatial feature delineation (i.e., structure and shape properties) the concept of urban structure types (USTs) has become a commonly used tool in urban planning. Specifically, USTs are defined by their functional and physical properties and through environmental features (HEIDEN et al. 2012:361). Functional properties mainly describe the land use, whereas physical properties are referring to e.g. the pattern of built-up areas or surface materials. In addition, environmental features consider climatic and hydrological aspects (PAULEIT & DUHME 2000:2f.).
BANZHAF & HÖFER (2008) delineated USTs by means of aerial images at a spatial resolution of 40 cm and Official Topographic-Cartographic Information System (ATKIS; German abbreviation for Amtliches Topographisch-Kartographisches Informationssystem) data of the Federal Republic of Germany. First, an object-based classification was conducted to detect different types of buildings and open spaces as well as their spatial relation. Next, the percentage of vegetation was determined using the normalized difference vegetation index (NDVI). Finally, the specific USTs have been determined accounting for road and railway networks, building properties, the percentage of impervious and vegetated surfaces, as well as thematic attributes from ATKIS.
A different approach to derive USTs was conducted by WALDE et al. (2014) using Quickbird imagery and an airborne light detection and ranging (LiDAR) data. Initially, a land cover classification with six information classes was derived and, similar to the above-mentioned approach, a road network has been employed for delineating single blocks

of USTs. Afterwards, the derived land cover classification and the blocks were used to generate neighborhood graphs. The classification of the USTs is based on selected graph measures, which were determined using the variable importance feature of the random forests classifier. WALDE et al. (2014:603) achieved an overall accuracy of 87 %, whereas the one obtained by BANZHAF & HÖFER (2008:137) was lower for a heterogeneous test site (i.e., 73 %) and higher for a homogenous test site (i.e., 93 %).

2.5.2 Urban environmental research

Urban environmental research mainly deals with the identification and monitoring of impacts caused by urbanization on the environment (MILLER & SMALL 2003:130). In this context, many studies focus on urban climate and the most of them address the urban heat island effect (CHEN et al. 2006, IMHOFF et al. 2010, TAM et al. 2015). In this framework, the land surface temperature (LST) is the most widely applied parameter and can be estimated by means of thermal infrared remote sensing sensors. CUI & SHI (2012:3ff.) analyzed the local environmental effects of urbanization in Shanghai by considering historical data on land use and population change, meteorological parameters, NDVI, and LST. The findings highlight that urbanization in Shanghai induced a massive decrease in agricultural land use due to urban expansion, while the urban heat island effect increased significantly. Furthermore, KAUFMANN et al. (2007) investigated the impact of urban growth on temporal and spatial patterns of precipitation in the Pearl River Delta region (which experienced a massive spatial expansion of urban areas) by means of Landsat imagery and meteorological data. Indeed, urbanization had an impact on the spatial and temporal pattern of precipitation and led to a decrease of precipitation locally.
Additionally, there are also studies focusing on the development of adaptation strategies to climate change for urban areas (CARTER et al. 2015, GILL et al. 2007, MASSON et al. 2014, HALLEGATTE 2009). For instance, GILL et al. (2007:116) assessed the effectiveness of green infrastructure in urban areas for climate change adaptation; indeed, their employment is becoming more and more important since they provide a cooling effect, hence reducing the magnitude of the urban heat island effect (CARTER et al. 2015:31).

2.5.3 Urban growth monitoring

In recent years, the assessment of urban growth using remote sensing data has been conducted in numerous studies. As illustrated in Figure 2.3 the number of published journal articles on urban growth mapping increased rapidly since the beginning of the last decade. In particular, the increase of publications coincides with the increase of VHR data availability and free access to data archives. Many studies are using Landsat scenes as basis for urban growth monitoring (BAGAN & YAMAGATA 2012, LI et al. 2015, QUAN et al. 2015, SCHNEIDER & WOODCOCK 2008, TAUBENBÖCK et al. 2012, VILLA 2012). LI et al. (2015:81ff.) employed a random forests classification on Landsat time-series between 1984 and 2013 for Beijing. The method comprises annual classification of urban areas

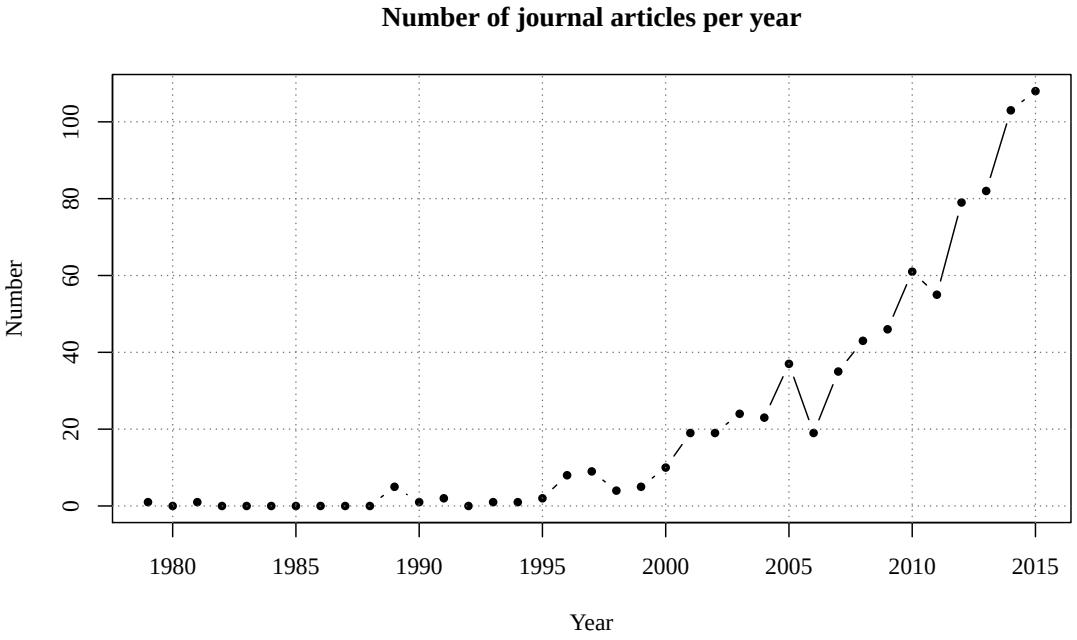

Figure 2.3: Number of published journal articles on urban growth mapping derived from a Scopus search on January 06, 2016 (keywords: urban growth remote sensing).

to precisely detect the changes, while possible misclassifications are excluded using a temporal consistency probability function, which compares the class label of the target year in a temporal domain. The results of this approach achieved an overall accuracy of more than 90 % for the urban land cover over different years. Moreover, VILLA (2012) proposed a classification approach on single date Landsat scenes to map the urban growth occurred in the Milan area for the years 1984, 1991, 1996, 1999, and 2003. Specifically, urban areas are discriminated from other land cover classes by thresholding soil and vegetation indices derived from Landsat's middle infrared and blue bands. The combined overall accuracy of the resulting maps is around 79 %.

So far, despite the proven effectiveness of SVM in a variety of classification problems, the number of studies where they have been employed in the context of urban growth mapping is very limited. A search on Scopus indicates only seven journal articles on this topic. Most of these studies used Landsat data and focused on a single study site (GRIFFITHS et al. 2010, MEGAHED et al. 2015, NEMMOUR & CHIBANI 2006, RIENOW & GOETZKE 2015, RÖDER et al. 2015). In addition to Landsat images, GRIFFITHS et al. (2010:429ff.) used synthetic aperture radar (SAR) data for the years 1990, 2000, and 2006. Further on, the authors included phenological information by means of bi-temporal Landsat images for each period. After the SVM classification post classification refinements were applied, hence enhancing the overall accuracy which resulted in the range between 83.5 and 84.7 %. XIAO et al. (2014) and PANDEY et al. (2013) employed a SVM based classification for urban growth monitoring by means of nighttime light data, Landsat images, and multi-source NDVI data, as well as nighttime light data in combination with SPOT vegetation

data, respectively.

Additionally, some studies focused on characterizing the spatial pattern and properties of growing urban areas. In this context, HEROLD et al. (2003b) applied spatial metrics on historical aerial and high resolution images to analyze the spatiotemporal growth pattern for Santa Barbara between 1930 and 2001. Also SCHNEIDER & WOODCOCK (2008) conducted a study analyzing the growth pattern of 25 cities across the globe including statistical census data in addition to remote sensing data and spatial metrics.

Table 2.1: Global urban extent and land cover maps based on Earth observation data sorted by their spatial resolution.

Name	Data basis	Definition of urban areas	Spatial resolution	Citation
Global Urban Footprint (GUF)*	TerraSAR-X, TanDEM-X	Built-up area with a vertical component	12 m	ESCH et al. (2012, 2013)
Global Human Settlement Layer (GHSL)*	Landsat	Built-up area	15–75 m	PESARESI et al. (2013, 2015)
GlobCover	ENVISAT MERIS	Artificial surfaces and associated areas (Urban areas >50 %)	309 m	ESA (2016)
MODIS 500 m	MODIS	Built-up area (>50 %), minimum mapping unit: 1 km²	463 m	SCHNEIDER et al. (2009)
MODIS 1 km	MODIS, nighttime lights, population density	Urban area	927 m	SCHNEIDER et al. (2003)
Global Impervious Surface Area (IMPSA)	nighttime lights, global population count data	Density of constructed impervious surface area	927 m	ELVIDGE et al. (2007)
Global Rural-Urban Mapping Project (GRUMP)	vector data, nighttime lights, census data	Urban extent	927 m	CIESIN (2011)
Global Land Cover (GLC2000)	SPOT, nighttime lights	Artificial surfaces and associated areas	988 m	BARTHOLOMÉ & BELWARD (2005)

**in progress*

2.5.4 Urban extent data on global scale

Previous sections highlighted the importance of outlining urban area extent. However, monitoring urbanization at the global scale remains a challenging task due to the heterogeneity of this specific information class and currently available global urban maps generally exhibit limited spatial resolution. Table 2.1 reports 8 currently available data sets, whereby two of them are still being processed and not yet publicly accessible. As one can notice, they have been derived from both optical and radar remote sensing imagery and, in some cases, also vector and census data have been employed. Six of the 8 datasets are binary urban extent maps; instead, the remaining two, i.e. the GlobCover and Global Land Cover, are land cover maps at a resolution of 309 m and 988 m, respectively. GlobCover is currently the accessible product with the highest spatial resolution and it is followed by the MODIS 500 m product. It is worth pointing out that there exist consistent differences between the urban extent between different maps. The Global Rural-Urban Mapping Project (GRUMP) map estimates a total urban area of 3,524 thousand km², whereas the other maps cover an area between 308 and 727 thousand km² (POTERE & SCHNEIDER 2009:271). Due to this high variation, the accuracy of each product should be questioned. SCHNEIDER et al. (2010:1743) conducted an accuracy assessment on the six currently available maps over 140 cities. The applied Cohen's Kappa coefficient (COHEN 1960) indicated the highest accuracy for MODIS 500 m with a mean Kappa value of 0.65, whereby the GRUMP map achieved the lowest accuracy with a mean Kappa of only 0.2.

The Global Human Settlement Layer (GHSL) and the GUF are not yet accessible, but an alpha versioan of the GHSL has been shared among the user community in October 2014 (PESARESI et al. 2015:840). The GHSL is developed by the Joint Research Center (JRC) of the European Commission, while the GUF is produced at the German Remote Sensing Data Center of the German Aerospace Center (DLR). As reported in Table 2.1 the GUF dataset is processed by means of TerraSAR-X add-on for digital elevation measurement (TanDEM-X) and TerraSAR-X, while the GHSL is based on optical data. The published alpha version of the GHSL is based on Landsat data, but it is intended to integrate Sentinel-1 and 2 data in future releases (PESARESI et al. 2015:840). Since the GUF is used in this study, a detailed description of the applied methodology is provided in the following paragraph.

First, it is worth noting that in the framework of the GUF a different definition of urban areas is considered with respect to that applied in this thesis. Indeed, the GUF solely identifies built-up areas (which have a vertical component) and excludes flat urban surfaces, such as roads or airport runways (ESCH et al. 2012:8). This is due to the fact, that the GUF is derived from SAR images, where the appearance of the illuminated target depends on its geometrical and dielectric properties. Here, buildings result in high values of the backscattered signal, while flat urban surfaces are associated with low values which are also typical of other information classes. The used TanDEM-X and TerraSAR-X images were recorded in stripmap mode at a spatial resolution of approximately three meters.

At the beginning of the GUF processing chain a texture layer, i.e. the so-called speckle divergence, is calculated based on the analysis of local speckle (which is a characteristic of coherent imaging sensors describing noise caused by several scatterers with different surface properties within a pixel) characteristics. Since urban areas are very heterogeneous, they generally result in high texture. The classification of urban pixels is performed automatically by means of one-class SVM (ESCH et al. 2013:1618f.). The GUF is produced at a spatial resolution of 12 m, however, a public domain version will be available at 84 m resolution (ESCH et al. 2012:6).

Besides the aforementioned global urban maps, recently published journal articles show that there are further attempts to derive the urban land use class globally by means of medium resolution data (CHEN et al. 2015, DUAN et al. 2015). CHEN et al. (2015:13f.) employed a land cover classification on different test sites located on five different continents using Landsat imagery. The objective of this study was to use the developed pixel- and object-based classification approach on a global scale to derive a land cover classification with ten classes, including the urban class. DUAN et al. (2015:2174ff.) aimed at generating a global urban map with a spatial resolution of 15 m using Phased Array L-band Synthetic Aperture Radar (PALSAR) and Advanced Spaceborne Thermal Emission and Reflection Radiometer (ASTER) data. This approach employs automatic collection of training samples and SVM-based classification. Moreover, this study tested the approach on 75 cities across the globe and the results achieved an overall accuracy of 84.4 % and a Kappa coefficient of 0.628.

3 Study areas

In this thesis, the urban extent was derived for 8 study areas for two time frames, namely 2002–2003 and 2013–2015. In particular, they correspond to the urban supersites selected by the SB-04 group of the GEO (WENG et al. 2014b:19). Specifically, the list include: Athens (Greece), Atlanta (USA), Beijing (China), Istanbul (Turkey), Los Angeles (USA), Mexico City (Mexico), Sao Paulo (Brazil) and Hong Kong (China). Concerning the last one, here the study site is actually extended to the entire Pearl River Delta, which also comprises the cities of Dongguan, Foshan, Guangzhou, and Shenzhen. The spatial distribution of the study areas is illustrated in Figure 3.1. The selected study sites cover urban areas of different sizes and patterns and in some cases a huge spatial expansion occurred between the investigated times.

Figure 3.1: Location of the study areas.

According to THE WORLD BANK GROUP (2015:21f.), the metropolitan area of the Pearl River Delta has recently overtaken Tokyo and it is now estimated to be the biggest megaregion on the Earth with 42 million inhabitants in 2010 and an urban area of approximately 7,000 km².

China's capital Beijing is the second urban area per population among the study areas. Since the 1980s it has witnessed an enormous spatial expansion and simultaneously the population increased from 8.72 million in 1987 to 19.62 million in 2010 (LI et al. 2013:3).

Istanbul is Turkey's most populated city and experienced a high population growth of 32 %

from 2000 to 2010 with over 13 million inhabitants. It is estimated that the population will surpass the 15 million mark in 2017 (TURKSTAT 2013).

The city Sao Paulo was home to 11,253,503 inhabitants in 2010 covering an area of 1,521 km², while predictions indicated a population of nearly 12 million in 2015 (BRAZILIAN INSTITUTE OF GEOGRAPHY AND STATISTICS 2014).

The population of central Mexico City increased only 1.3 % between 1990 and 2000. At present more than 8 million people are residing in Mexico City, but approximately 20 million people live in the metropolitan area (MERLÍN-URIBE et al. 2013:399).

The population of Los Angeles grew of about 2.7 % between 2000 and 2010, when the number of inhabitants was estimated around 3,79 million (U.S. CENSUS BUREAU 2012).

The population of Athens rose by 3.5 % from 2,89 million inhabitants in 2004 to 2,99 million in 2009 (EUROSTAT 2015).

The study area with the lowest population is Atlanta with 456,002 inhabitants in 2014 (U.S. CENSUS BUREAU 2012); nevertheless its increase with respect to the year 2000 was of 8.2 %.

4 Methods

The developed methodology aims at automatically delineating the urban area extent and, hence, is suitable for characterizing urbanization over time. In particular, the implemented chain includes, for a given study area, the acquisition and pre-processing of available Landsat imagery, the enhancement of the corresponding GUF, the classification of urban areas, and the corresponding accuracy assessment. A block scheme of the proposed approach is shown in Figure 4.1.

Different spectral indices are generated from Landsat data and, afterwards, key temporal statistics are derived for each of them. These are then employed first for enhancing the GUF over the investigated region (from which training samples are extracted based on specific rulesets) and, then, for identifying urban areas by means of SVM (a simple DT classifier has also been tested for comparison purposes). The image processing and classification was mainly performed using gdal tools (Version 1.11.3, Open Source Geospatial Foundation), while the analysis and data visualization has been carried out with R Statistics (Version

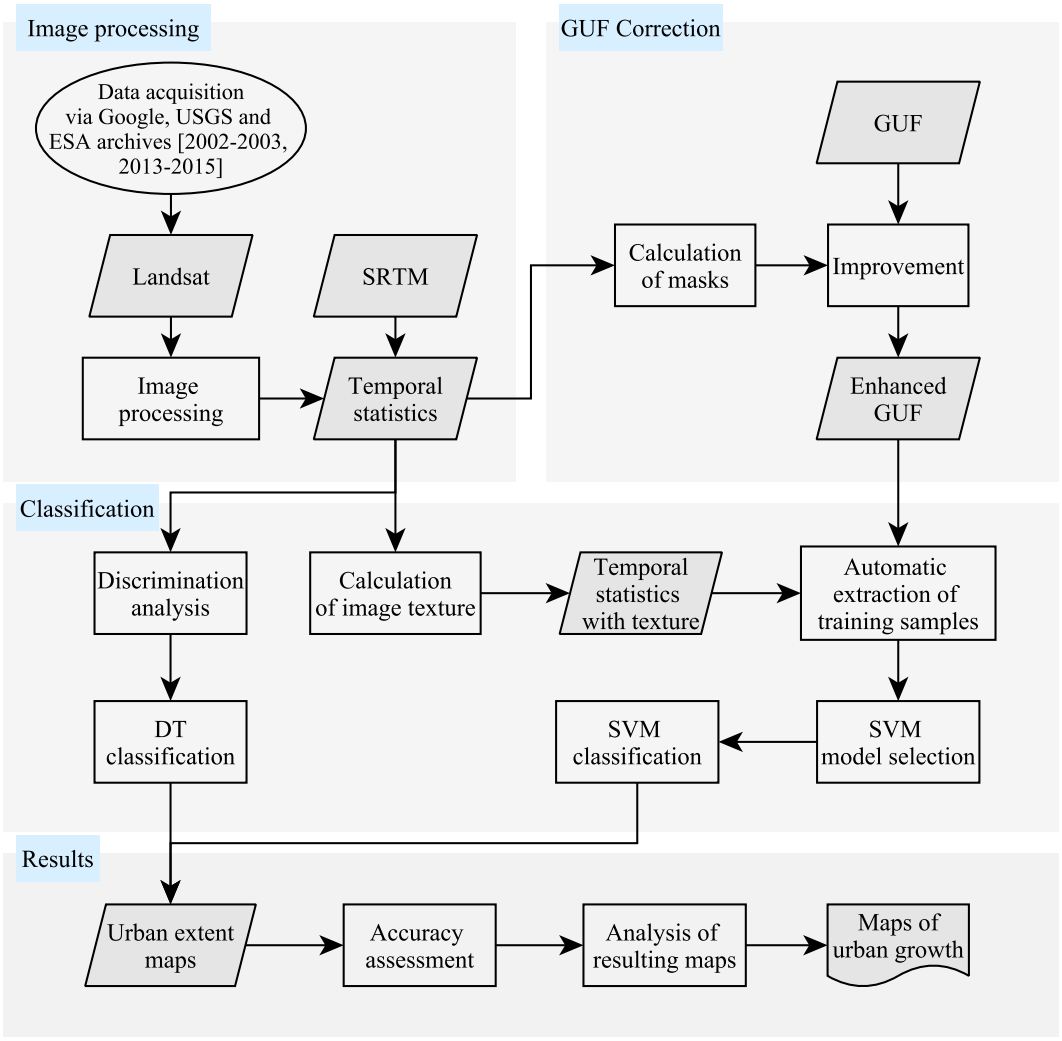

Figure 4.1: Block scheme outlining the proposed methodology to delineate urban area extent.

3.2.2, R Foundation for Statistical Computing, Vienna, Austria) and ArcGIS (Version 10.3.1, ESRI, Redlands, USA). In the following sections a description of the presented technique is given into detail.

4.1 Data basis

A brief description of the employed remote sensing data along with their purpose is reported in Table 4.1. The Landsat images have been downloaded from the databases of Google Earth Engine, ESA, and the United States Geological Survey (USGS), whereas the GUF and the Shuttle Radar Topography Mission (SRTM) elevation data have been provided by DLR.

Table 4.1: Data used in the presented study.

Name	Spatial resolution	Intended use	Time period
Landsat imagery	30 m	GUF enhancement and automatic classification of urban areas	2002–2003, 2013–2015
Global Urban Footprint (GUF)	12 m	Collection of training samples	2011–2013
Shuttle Radar Topography Mission (SRTM) data	1 arc-second (~30 m)	Classification of urban areas and improvement of the GUF	2000
Very high resolution imagery (Google Earth)	<2.5 m	Validation of derived urban extent maps	2002–2003, 2013–2015

4.1.1 Landsat data

In this study, Landsat-5/7/8 scenes from TM, ETM, and OLI sensors, respectively, have been used for the two investigated time periods, namely 2002–2003 and 2013–2015. In

Table 4.2: Number of total Landsat scenes which were included in the calculation of the temporal statistics for each study area.

Study area	Area [km²]	2002-2003 (TM/ETM)	2013-2015 (OLI)	Path and Row (WRS-2)
Athens	11,209	59 (23/36)	96	182034, 183033, 183034
Atlanta	15,053	62 (45/17)	49	019036, 019037
Beijing	32,230	121 (43/78)	118	122032, 122033, 123032, 123033, 124032
Istanbul	16,564	93 (33/60)	88	179031, 179032, 180031, 180032
Los Angeles	25,401	86 (31/55)	148	040036, 040037, 041036, 041037
Mexico City	10,852	31 (5/26)	93	026046, 026047, 027046
Pearl River Delta	42,761	130 (68/62)	81	121044, 121045, 122044, 122045, 123044, 123045
Sao Paulo	14,213	85 (37/48)	94	218076, 218077, 219076, 219077

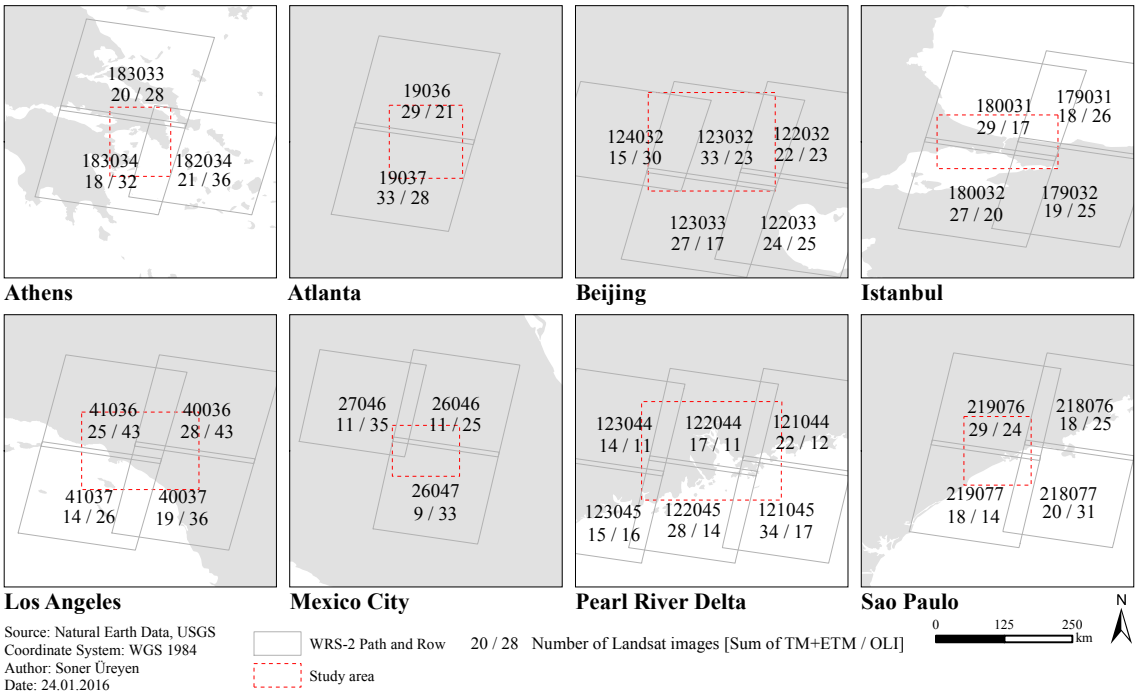

Figure 4.2: Depiction of the number of Landsat images for each path-row combination and the extent of the study areas.

particular, Level 1T data with standard terrain correction have been considered. Here, it is worth noting that, due to the limited number of scenes available in 2002–2003 for the Pearl River Delta site, the corresponding analysis has been extended to the year 2001.
Only Landsat scenes intersecting the different areas of interest (AOI) with cloud cover lower than 30 % have been employed. The corresponding path-row combinations according to the Worldwide Reference System 2 (WRS-2) are listed in Table 4.2 and depicted in Figure 4.2. The varying number of images among the study areas depends on both their location and size (indeed, the most of the archived Landsat data covers North America, Australia, and Eastern China (WULDER et al. 2015:4)). The highest availability occurs for Los Angeles in the period 2013–2015 (i.e., 148); the lowest for Mexico City in the period 2002–2003 (i.e., 31). In some cases the data availability for different path-row combinations referring to the same study site is rather heterogeneous.

4.1.2 The Global Urban Footprint

It has been pointed out in the previous chapter, that the GUF is a binary mask globally delineating built-up areas at 12 m spatial resolution. As discussed in ESCH et al. (2013) the corresponding GUF processor initially generates 7 classification maps based on as many thresholds on the original radar backscattering amplitude and the corresponding speckle divergence texture feature. The areas marked as urban increase from version 1 (which tends to underestimate the actual built-up extent) to version 7 (which tends to overestimate the actual built-up extent). An automatic approach is then applied to identify the optimal version. An example of the 7 GUF versions for the area of Athens is reported in Figure 4.3.

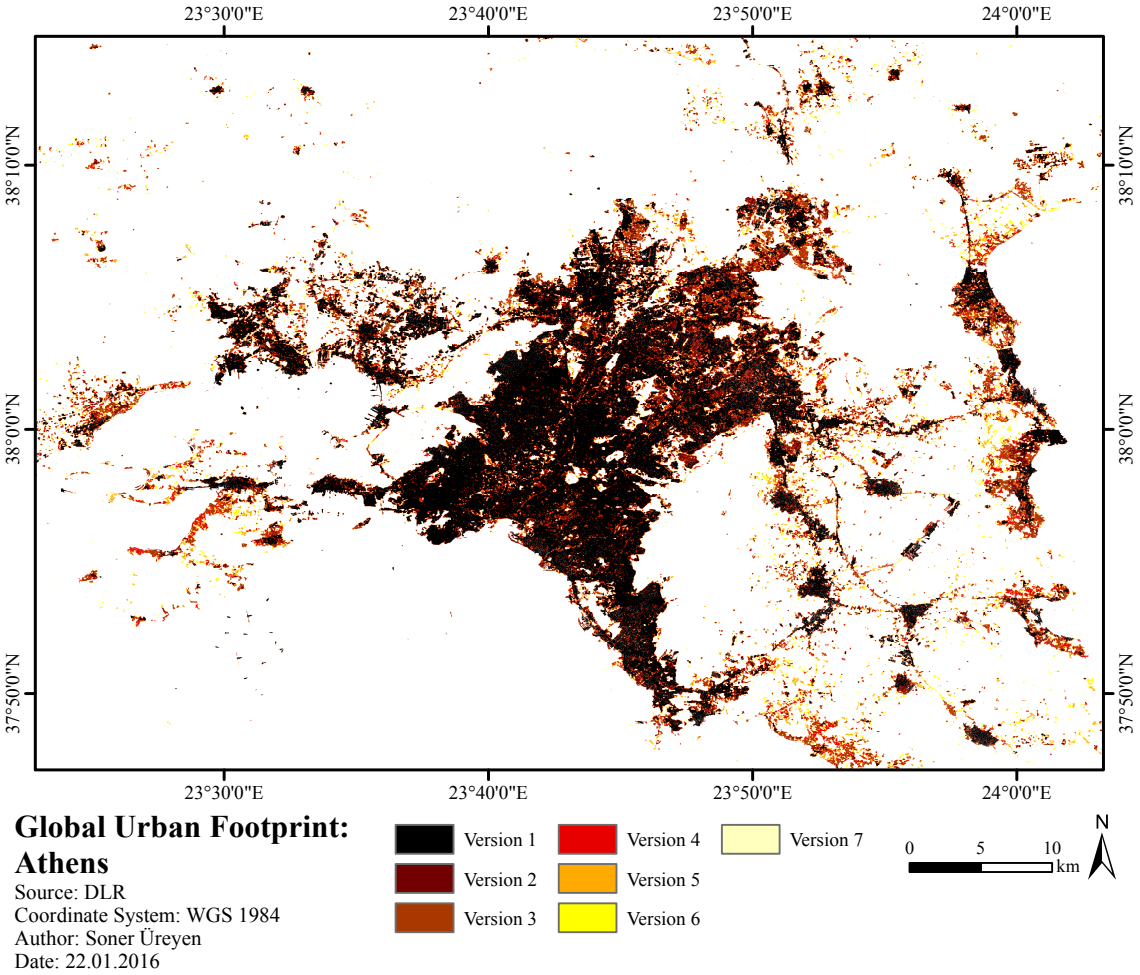

Figure 4.3: Different GUF versions for the investigated Athens study area.

4.1.3 Elevation data

SRTM is currently the mostly used DEM with coverage between 60° north and 54° south latitude (U.S. GEOLOGICAL SURVEY 2014). It is available both at a spatial resolution of 3 and 1 arc-second (which corresponds at the equator to 90 and 30 m, respectively) and it can be downloaded in 1x1° tiles. 9 tiles were necessary to cover the entire extent of the Pearl River Delta, while only two are needed for the Mexico City site. The SRTM data are provided as archives where the raster files are stored in Sun Raster (.sras) file format. After downloading and unpacking the data, header files are generated for each SRTM tile and afterwards the tool gdal_merge.py has been used to mosaic them.

4.2 Image processing

This section describes the pre-processing of the Landsat data, as well as the selection of the corresponding spectral indices and temporal statistics derived for performing the automatic classification of urban areas. Additional texture features used for improving the results in low-density areas are also explained.

4.2.1 Processing of Landsat data

The processing of the downloaded Landsat images starts with the conversion of the digital numbers (DN) to at-sensor radiance. Indeed, the Landsat sensors record the electromagnetic radiation, convert it to radiance, and then scale it to DN in 8-bit integer format for the TM and ETM sensors and 16-bit integer format for the OLI sensor. To perform the conversion from DN to radiance values Equation (4.1) is used:

$$L_\lambda = gain_\lambda * QCAL_\lambda + bias_\lambda \tag{4.1}$$

where, for the generic spectral band λ, L denotes the at-sensor radiance in units of watts per meter squared per steradian per micrometer $(W/(m^2 * sr * \mu m))$, while $gain_\lambda$ and $bias_\lambda$ are band specific rescaling factors delivered in the image meta data file, and $QCAL_\lambda$ is the quantized calibrated pixel value in DN (U.S. GEOLOGICAL SURVEY 2015b).
It is worth pointing out that in this study no atmospheric correction has been applied to the Landsat Level 1T data, since, as described in SONG et al. (2001:232), this is not mandatory when performing land cover classification with training data collected from the same image to be classified.
After the calibration, cloud and cloud shadow masks are generated for each Landsat image to exclude these areas from the analysis. To this purpose the Function of mask (Fmask) algorithm version 3.2 is applied, which has been recently updated and is widely used for automatic cloud detection (ZHU et al. 2015:272). Besides pixels covered by clouds and cloud shadows, the algorithm also identifies snow, clear land and clear water pixels and it can be applied to any Landsat image. These features are extracted by means of a ruleset based on the NDVI, the Normalized Difference Snow Index, and the Brightness Temperature (ZHU & WOODCOCK 2012:84f.).
Next, spectral indices are calculated for each single Landsat image. In this context, to ensure an effective delineation of the urban areas, a thorough analysis has been carried out to identify the most suitable set. The final list is reported in Table 4.3. Empirical evidence indicates the applicability of the Normalized Difference Built-Up Index (NDBI), Modified Normalized Difference Water Index (MNDWI), and Normalized Difference Vegetation Index (NDVI) to extract several land cover classes.
The NDBI has been proposed by ZHA et al. (2003) and applied to extract built-up areas in many studies (CHEN et al. 2006, XU 2007); nevertheless, due to the use of the middle infrared band (i.e., TM/ETM 5 and OLI 6) this index is also influenced by dry vegetation, since, in the presence of vegetation with low water content the reflectance in this band increases (XU et al. 2013:1441). This results in NDBI values comparable to those of urban areas. The NDVI and Normalized Difference Middle Infrared (NDMIR) are applied to overcome this problem.
The NDMIR uses both middle infrared bands, thus being sensitive to vegetation moisture (LU et al. 2004). The NDVI was first published by ROUSE et al. (1973) and since then it

has been widely employed in a variety of land cover applications as well as in the context of urban extent classifications (MASEK et al. 2000, SCHNEIDER 2012). Moreover, the MNDWI is used to discriminate water from urban areas. XU (2006:3027) enhanced the performance of the NDWI by replacing the NIR with the MIR band, which leads to a reduction of noise from built-up areas.
In addition to the previous indices, two further spectral normalized ratios have been derived based on the findings of ZHOU et al. (2014) to improve the discrimination between urban areas and bare soil/bare rocks. In particular, the blue band has been used as it proved quite effective in separating these two information classes (U.S. GEOLOGICAL SURVEY 2015a). The resulting novel indices are referred to as Normalized Difference Red Blue (NDRB) and Normalized Difference Green Blue (NDGB) and an analysis of their effectiveness is outlined in the following pages.
Overall, the 6 indices have been derived for a total of 1,434 scenes by means of a dedicated tool developed at the German Remote Sensing Data Center of the DLR in the Land Surface Department.

Table 4.3: Spectral indices used in the presented approach.

Spectral index	Formula	Citation
Normalized Difference Built-Up Index (NDBI)	(MIR1-NIR)/(MIR1+NIR)	ZHA et al. (2003)
Modified Normalized Difference Water Index (MNDWI)	(Green-NIR)/(Green+NIR)	XU (2006)
Normalized Difference Vegetation Index (NDVI)	(NIR-Red)/(NIR+Red)	ROUSE et al. (1973)
Normalized Difference Middle Infrared (NDMIR)	(MIR1-MIR2)/(MIR1+MIR2)	LU et al. (2004)
Normalized Difference Red Blue (NDRB)	(Red-Blue)/(Red+Blue)	ZHOU et al. (2014)
Normalized Difference Green Blue (NDGB)	(Green-Blue)/(Green+Blue)	ZHOU et al. (2014)

4.2.2 Derivation of Landsat based temporal statistics

To characterize the different behavior over time of the urban class with respect to all the others, key temporal statistics are extracted for the abovementioned 6 spectral indices. In particular, for the given study site all the indices derived for each available scene are first sorted in chronological order from the oldest to the newest. Then the temporal maximum, minimum, mean, standard deviation and mean slope are calculated.
Specifically, the temporal maximum and minimum corresponds to the observed maximum and minimum values, respectively, of the entire multi-temporal sequence. Instead, the temporal mean and standard deviation characterize the average and the variation over time,

respectively. The mean slope is derived using following Equation (4.2):

$$mean\ slope = \frac{1}{n-1} \sum_{t=1}^{n-1} \left(\left| \rho_i^t - \rho_i^{t+1} \right| \right) \tag{4.2}$$

where n denotes the number of scenes, t the acquisition date, and ρ the pixel value. This measure allows to enhance the discrimination between land cover classes which are more stable through the entire multi-temporal sequence (e.g., urban areas, bare soil, bare rocks) with respect to those that exhibit a higher variability (e.g., vegetated areas). In addition to the resulting 30 temporal indices, an additional feature is considered reporting the number of valid scenes (i.e., cloud-free) per pixel.

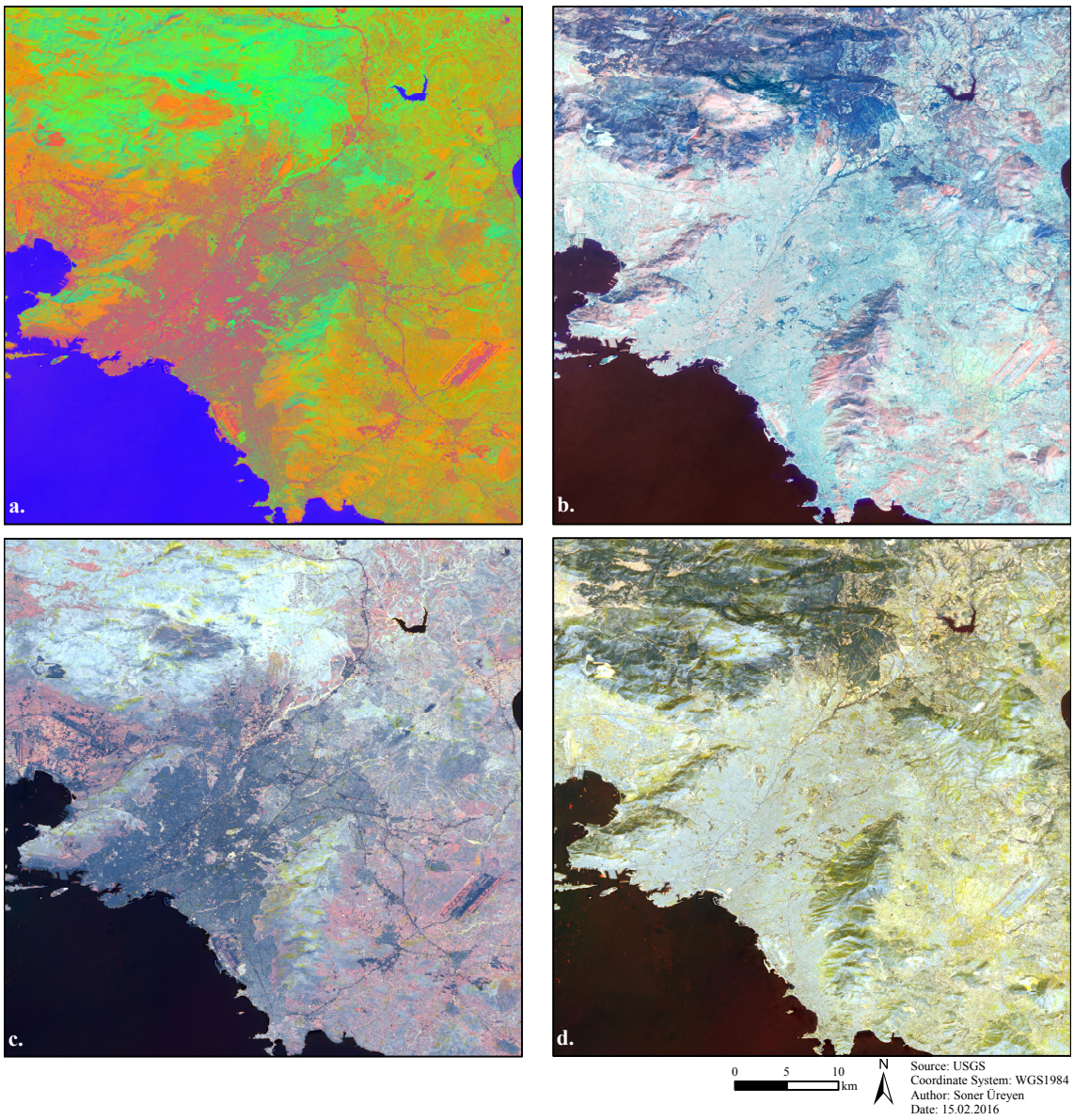

Figure 4.4: Temporal statistics (2013–2015) for Athens showing several false color composites (a.) Temporal Mean NDBI (Red), NDVI (Green), and MNDWI (Blue); b.) Temporal Mean NDBI (Red), NDRG (Green), and NDBG (Blue); c.) Temporal Maximum NDVI (Red), Mean NDVI (Green), and Minimum NDVI (Blue); d.) Temporal Maximum NDRG (Red), Mean NDRG (Green), and Minimum NDRG (Blue)).

Figure 4.4 demonstrates the Landsat based temporal statistics with different false color compositions. As one can see, the arrangement of the temporal mean NDBI, NDVI, and MNDWI makes it possible to visually differentiate between several land cover classes, while red tones correspond to sealed surfaces, green color reveals vegetated surfaces, and bright orange tones characterize fallow land as well as bare soil. The selection of the temporal indices maximum, mean, and minimum NDRG is very useful to highlight extraction sites which appear almost white or very light yellow.

To prove the effectiveness of the temporal indices in discriminating different land cover classes, for the 2013–2015 temporal mean of each spectral index in Figure 4.5 boxplots are reported showing the distribution of different information classes (for which a number of samples have been manually delineated by photo interpretation of Google Earth imagery for all the 8 investigated study sites). These include urban, water, vegetation, and barren land (composed of construction sites, mineral extraction sites as well as fallow land).

The temporal mean NDBI values for the classes urban, vegetation, and barren land are similar, but the interquartile range of the boxplots have little overlaps. The temporal mean NDVI and NDMIR are suitable to distinguish urban and vegetated areas, whereas the temporal mean NDGB and NDRB prove useful for separating urban areas from barren land (indeed, the boxplots have minor overlaps just among outliers). In addition, water surfaces are clearly separable from all the other land cover classes by means of the temporal mean MNDWI.

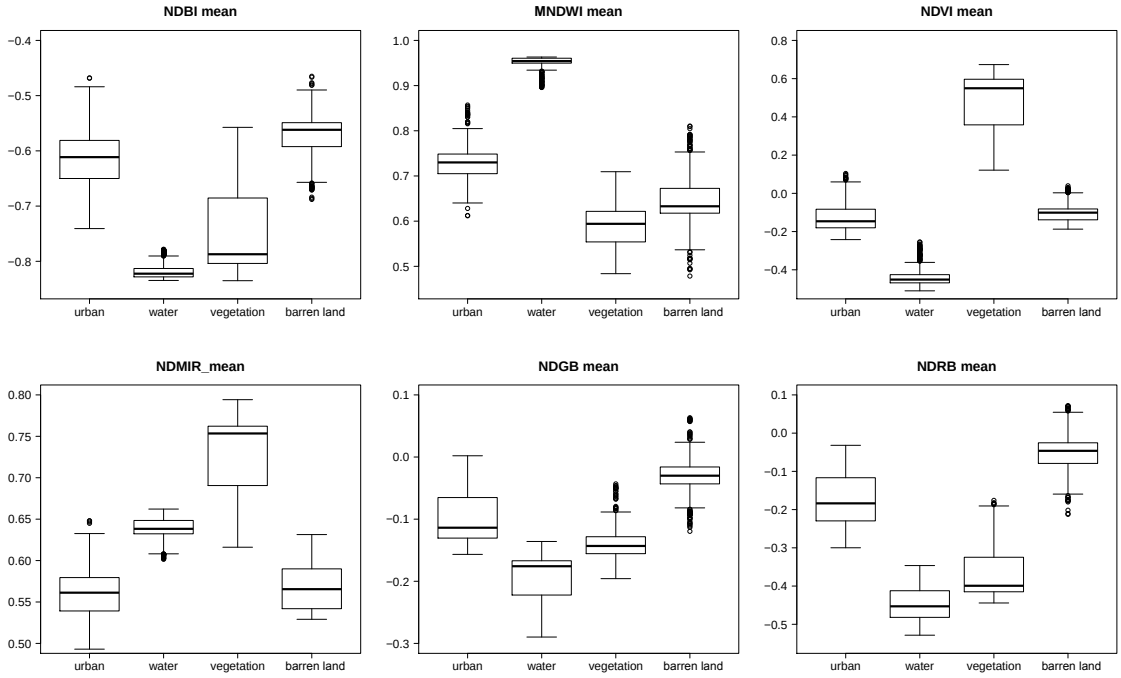

Figure 4.5: Boxplots demonstrating the applicability of the six indices with samples extracted from the temporal statistics (2013–2015) of all study areas.

4.2.2.1 Extraction of texture features

In order to improve the detection of suburban areas (which tend to be underestimated due to the 30 m spatial resolution of Landsat data), additional texture features have been extracted. Such areas are mostly characterized by a low share of built-up areas and a high share of vegetation, thus resulting in a heterogeneous environment compared to denser built-up areas. Accordingly, it is expected that they exhibit higher texture values.
By means of the pkfilter tool the local spatial mean and variance are extracted for the temporal mean of each of the considered 6 spectral indices. Here, since the filtering window size has a great impact, different values have been tested, namely 3x3, 5x5, 7x7, and 9x9 pixels. Afterwards, the corresponding coefficient of variation (CV) and the index

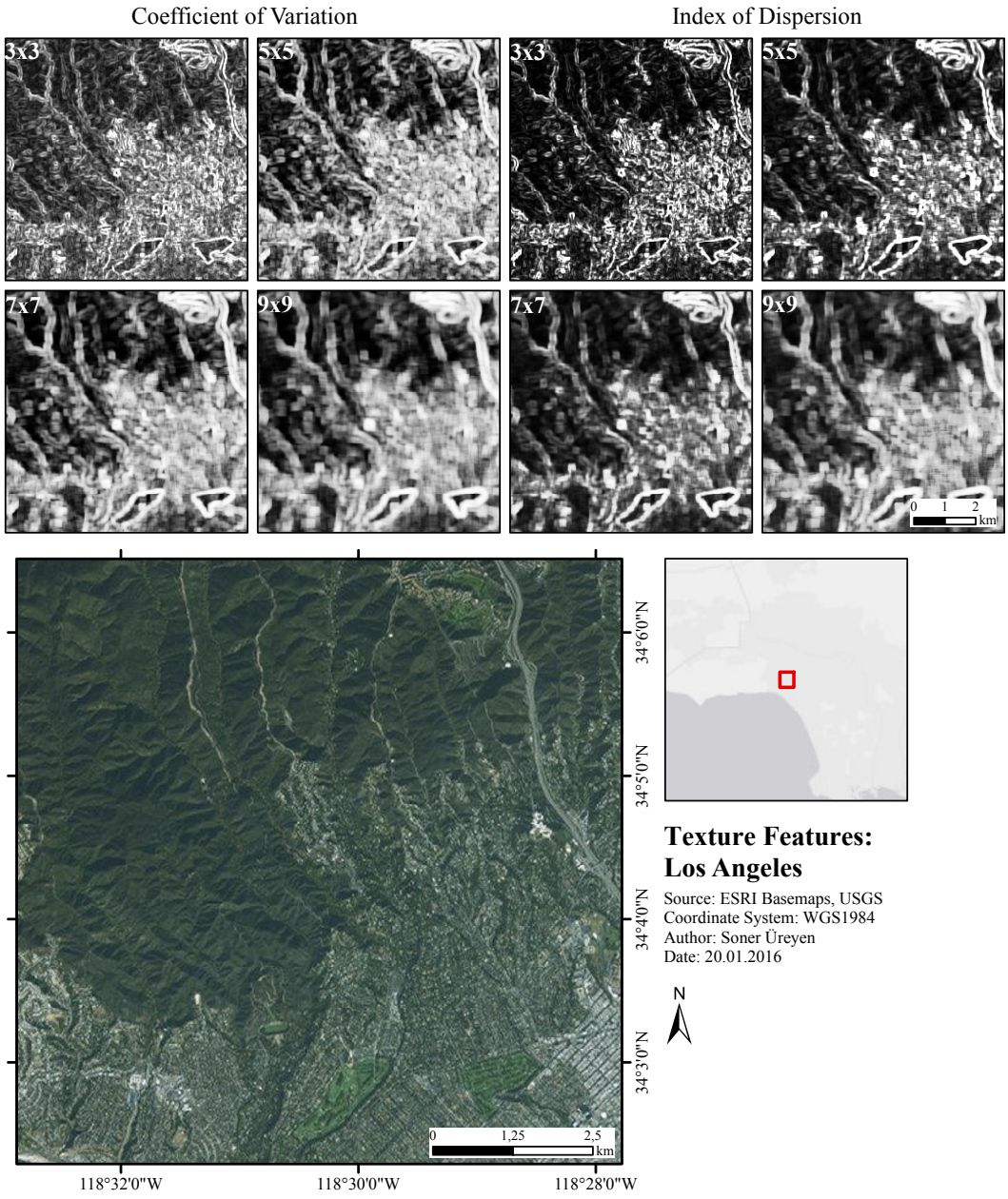

Figure 4.6: Comparison of computed texture features from the temporal statistics with different window sizes for the study area Los Angeles. The illustrated texture images are based on the temporal mean NDMIR.

of dispersion (ID, also known as the variance to mean ratio), have been computed. In details, the CV is calculated as:

$$CV = \frac{\sqrt{\sigma^2}}{\mu} \tag{4.3}$$

where $\sqrt{\sigma^2}$ is the local spatial standard deviation and μ the local spatial mean of a given pixel, while the ID is derived as:

$$ID = \frac{\sigma^2}{\mu} \tag{4.4}$$

with σ^2 representing the local spatial variance. Both CV and ID are often applied in remote sensing applications to highlight image heterogeneity (CULBERT et al. 2009:302; DEVRIES et al. 2007:416; WAI YEUNG YAN & SHAKER 2014:7665).

Figure 4.6 depicts the obtained texture features for the temporal mean NDMIR index with different window sizes. One can notice that, with increasing filtering size, the averaging effect strongly intensifies. Hence, being the scope of employing texture features to improve the classification of suburban areas where the target objects are mainly buildings with small extents, a size of 3x3 pixels has been finally chosen.

4.2.2.2 Analysis of the Digital Elevation Model

As mentioned previously, the main purpose of using the SRTM data is the enhancement of the GUF, which in some cases exhibits some overestimation since areas with complex topography can be misclassified as urban due to the corresponding very high backscattering. To this purpose, after resampling the SRTM to the Landsat spatial resolution using a bilinear interpolation, two terrain indices have been computed by means of the gdaldem tool, namely roughness and slope.

The roughness is defined as the difference between the maximum and minimum value among the given pixel and its surrounding 8 neighbors in a 3x3 window, whereas the slope is given by the maximum difference between the given pixel and its 8 neighbors in degrees

Table 4.4: Complete feature space which is used for the automatic classification of urban area extent.

Band	**Spectral indices (max., min., mean, stddev., mean slope)**	**Band**	**Texture (CV, ID)**	**Band**	**Terrain indices**
1–5	NDBI	32;33	NDBI	44	Roughness
6–10	MNDWI	34;35	MNDWI	45	Slope
11–15	NDVI	36;37	NDVI		
16–20	NDMIR	38;39	NDMIR		
21–25	NDRG	40;41	NDRG		
26–30	NDBG	42;43	NDBG		
31	Number of available scenes per pixel				

(WILSON et al. 2007:12). In some cases, roughness and slope are strongly correlated as for the Los Angeles study area where the corresponding correlation coefficient is equal to 0.98; nonetheless, they have been both taken into account while performing the final classification of the urban area extent. Accordingly, overall 45 features are finally taken into account, of which a summary is reported in Table 4.4.

4.2.3 Retrieval of water and vegetation masks

The 2013–2015 Landsat based temporal statistics have been used to generate additional masks for improving the GUF, mostly for removing trees wrongly categorized as urban due to specific characteristics of the investigated area, as well as ships present in the AOI at the time when the original radar imagery has been acquired. To this aim, by analyzing the boxplots in Figure 4.4, specific thresholds have been defined and tested for the most suitable indices. On the one hand a vegetation mask has been derived from the temporal mean NDVI and NDMIR (pixels showing values between 0.35 and 1 for the former, and between 0.7 and 1 for the latter have been marked as vegetation). On the other hand, a water mask has been calculated from the temporal mean and maximum MNDWI (pixels showing values greater than 0.9 for the former, and greater than 0.85 for the latter have been marked as water). Moreover, the temporal mean NDRB has been employed for excluding excavation or construction sites included in the GUF layer in those cases where they have a vertical steep component (pixel showing values between 0 and 1 have been marked as barren land). All the 3 masks have been computed by means of the gdal_calc.py tool and a qualitative assessment has been carried out by overlaying them to Google Earth (Pro Version 7.1.5.1557, Google Inc., Mountain View, California, USA) VHR imagery.

4.3 Enhancement of the Global Urban Footprint

The GUF layer has been employed in this study for automatically collecting training points for the periods 2002–2003 and 2013–2015, respectively. To this aim, it has been first resampled to the 30 m resolution of Landsat data (from its original 12 m resolution) and, to guarantee a perfect alignment, the pixels of the resized GUF have been snapped to those of the Landsat based temporal statistics. As already discussed, to ensure a collection of reliable training points it was necessary to minimize the amount of misclassified pixels in the GUF. Figure 4.7 reports some examples of the most common errors that had to be eliminated. Here, GUF version 7 (which is generally overestimating urban areas) is compared with imagery basemaps in ArcGIS indicating common sources of error; here illuminated targets such as ships or trees in close distance to water are characterized by high backscatter values such as built-up areas. The improvement has been applied to all 7 GUF versions using the gdal_calc.py tool and marking as non-urban all those GUF pixels originally labeled as urban which intersected any of the three masks described in Section 4.2.3. Furthermore, to exclude areas with complex topography, also pixels whose slope was higher than 25° have been always set as non-urban.

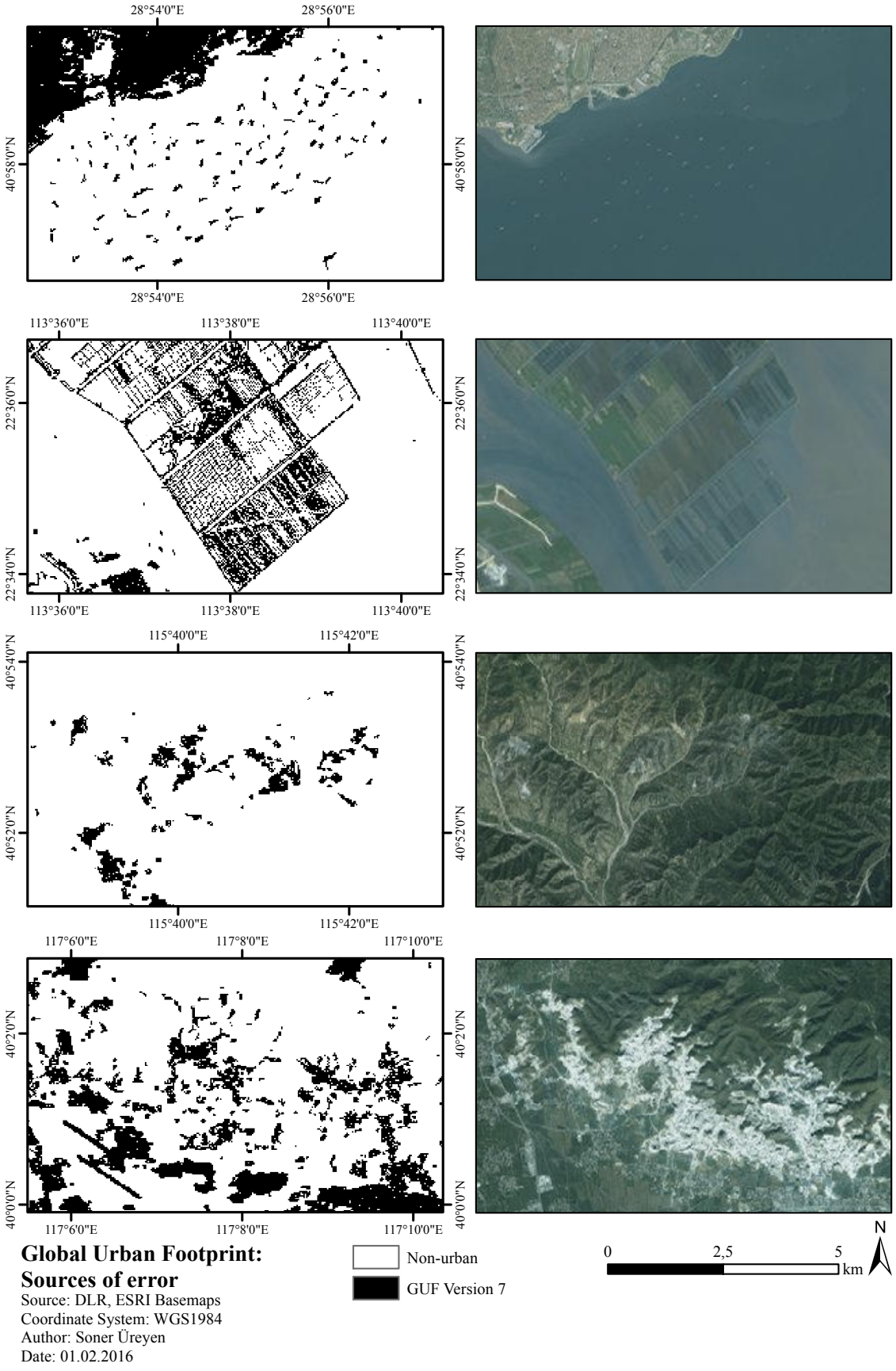

Figure 4.7: Depiction of the GUFs main sources of error at a spatial resolution of 30 m. 1. Ships classified as urban south of Istanbul, 2. Misclassification of aquaculture in the Pearl River Delta, 3. Effects of topography northeast of Beijing, 4. Excavation sites east of Beijing as a source of error.

4.4 Classification

This section presents the classification approaches employed for automatically outlining urban areas from the set of 45 selected features. On the one side, a linear classifier was applied for testing purposes to the 2013–2015 data; on the other side, a SVM classifier was used to delineate the urban extent for both the considered time frames.

4.4.1 Unsupervised DT classification

An unsupervised DT classification was first conducted to examine whether the defined classification problem is linearly solvable. For this reason, this has been solely tested for the 2013–2015 period. An extensive study has been carried out for all the 8 selected sites with the scope to derive a single ruleset which can be effectively applied to any study area. In particular, here only the temporal mean and standard deviation of the 6 chosen spectral indices have been taken into account and samples for the classes urban, water, vegetation, and barren land have been collected and separately analyzed for each AOI. The results of this activity are reported in Appendix A, while in Table 4.5 the final selected subset of 8 features along with the corresponding thresholds are described. To prove the relevance of accounting for the temporal standard deviation, in Figure 4.8 an example is given where it is clear the different behavior of urban classes with respect to agricultural fields showing a much higher variability over time. The final ruleset has been applied using the gdal_calc.py tool.

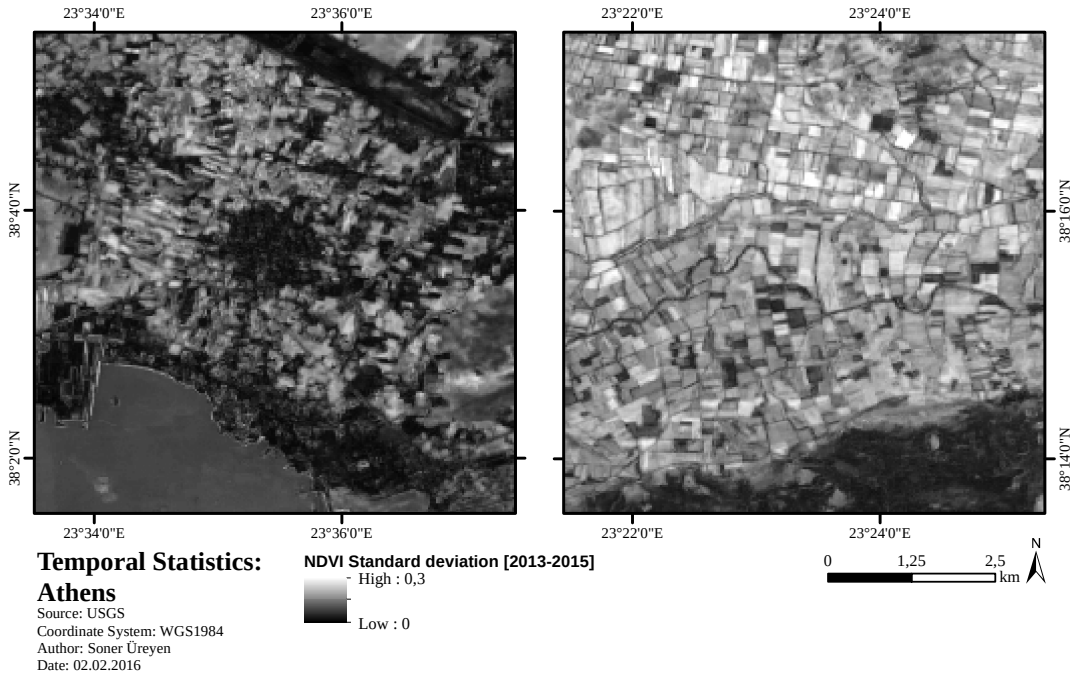

Figure 4.8: Feature NDVI standard deviation showing built-up areas on the left and agricultural areas and grassland on the right.

Table 4.5: List of employed features and corresponding thresholds to determine urban extent.

Feature	Thresholds
NDBI mean	-0.8 – -0.4
MNDWI mean	0.5 – 0.85
NDVI mean	-0.35 – 0.2
NDMIR mean	0.5 – 0.75
NDRB, NDGB mean	-0.5 – 0
NDBI, MNDWI, NDVI standard deviation	0 – 0.1

4.4.2 Proposed SVM based classification scheme

The block scheme of the proposed classification system based on SVM is shown in Figure 4.9 and it has been implemented as a shell script to be run in a Linux environment. The system is fully automated and solely requires as input the Landsat based temporal statistics for the selected indices together with version 1 and version 7 of the GUF for the area under investigation.

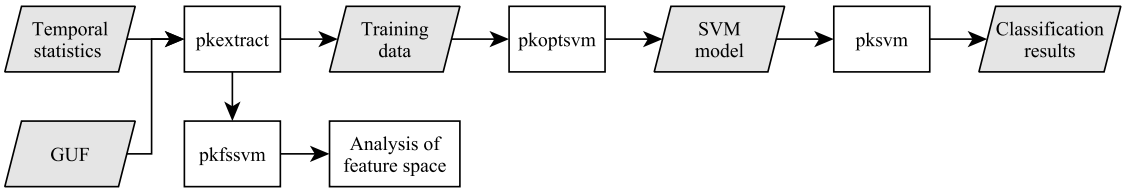

Figure 4.9: Block scheme of the proposed SVM based classification system.

4.4.2.1 Collection of training samples

The rationale of using the GUF, which has been produced with data acquired between 2011 and 2013, for automatically deriving training points for the proposed classification system is based on the assumption that urban growth has occurred rather than shrinkage over time. Then, specific criteria have been used for the two investigated time frames by properly using GUF version 1 (i.e., the one derived with the stricter thresholds on the original backscattering and speckle divergence, hence generally underestimating the actual built-up extent) and version 7 (i.e., the one derived with the softer thresholds on the original backscattering and speckle divergence hence generally overestimating the actual built-up extent).

When considering the 2013–2015 period, training samples for the urban class are randomly extracted from GUF version 1, since they are reasonably expected to remain urban (to exclude mixed pixels at the borders the pkfilter tool with the option `-erode` and a window size of 3x3 was also applied, while the gdal_sieve.py tool has been used to remove isolated settlements of size smaller than 3 pixels). Moreover, it is more likely that urban growth occurred at the fringes of GUF rather than appearance of new urban structures in areas far

away from already existing settlements. Accordingly, training points for the non-urban class are taken according with GUF version 7 except from a buffer along the outlines of the areas marked as urban where it is expected that urban growth took place (the pkfilter tool has been used with the option `-dilate` and a window size of 21x21 pixels, corresponding approximately to 315 m).

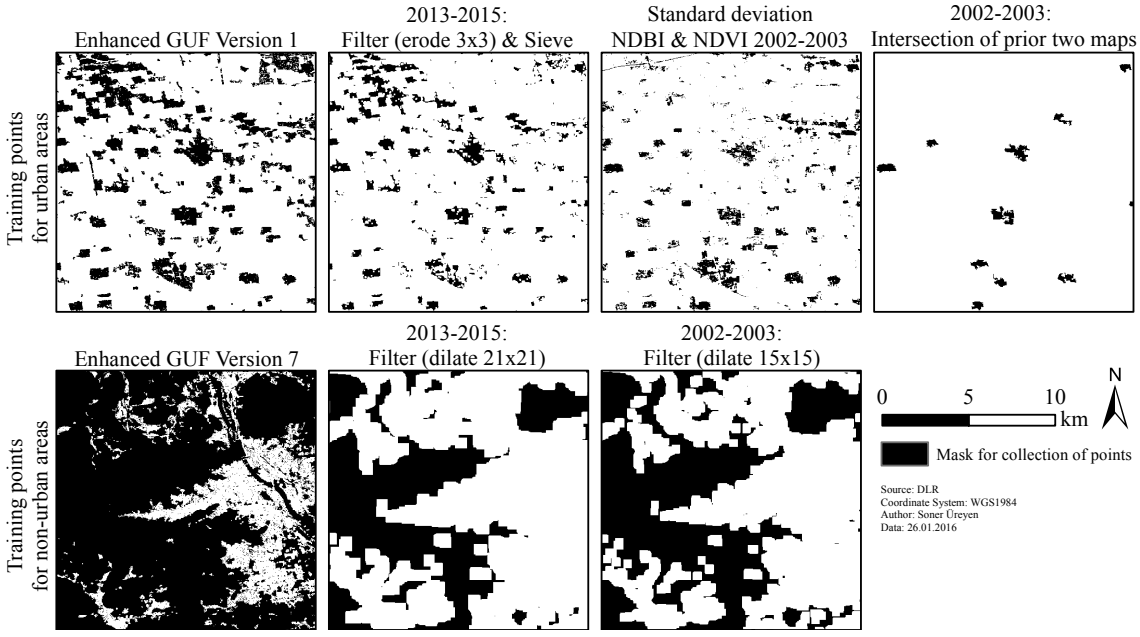

Figure 4.10: Computed masks for the collection of training samples for the urban and non-urban class.

When analyzing the 2002–2003 period, training samples for the non-urban class are randomly extracted according with GUF version 7, since – on the abovementioned assumption of urban growth over time – they are reasonably expected to be non-urban also before the time which the GUF refers to (however, to avoid problems that might arise in the case even GUF version 7 is underestimating the actual urban extent, a buffer has been created using the pkfilter tool with the option `-dilate` and a window size of 15x15 pixels). Instead, training samples for the urban class are derived from GUF version 1; nevertheless, to assess their reliability and exclude points that were actually not urban in 2002–2003 the corresponding temporal standard deviation of the NDBI and NDVI indices is analyzed. In particular, samples exhibiting values greater than 0.05 and 0.1, respectively, are rejected due to their high temporal variability in the considered time frame which is not typical of urban settlements.

As there exist no criteria for determining a priori the optimal amount of labeled points to generate, 9 different configurations have been tested with different numbers of training points. The complete list is reported in Table 4.6. Specifically, the idea behind these configurations is to investigate different proportions between the number of training points for both classes; hence training sets with a balanced and unbalanced relation between the two

information classes have been specified.
However, since results might vary depending on the specific selected training points, as a means for further improving the final performances and obtain a more robust classification map, for each considered criterion 10 different training sets are randomly created per each test site and time period resulting in as many corresponding urban extent maps. Then, assuming that errors are not correlated, a majority voting approach is applied (KITTLER et al. 1996:897; RONG YAN et al. 2003:23) where each pixel is associated with the urban class only in the case it is labeled as urban in at least 6 over 10 maps.

Table 4.6: Criteria tested to generate the training points for the proposed classification system.

		# of training points	
Scenario	**Relation**	**Urban**	**Non-urban**
1	unbalanced: one third	500	1,500
2		1,000	3,000
3	balanced	500	500
4		1,000	1,000
5	unbalanced: one half	500	1,000
6		1,000	2,000
7	unbalanced: one quarter	500	2,000
8		1,000	4,000
9	unbalanced	1 % of mask	1,000

4.4.2.2 Model selection and classification

After generating all the training sets, the pkoptsvm tool is employed to identify for each of them the optimal values for the learning parameters (i.e., the ones expected to provide the best possible discrimination between the urban and non-urban classes), namely C and γ since SVM with RBF kernels have been used. To this purpose a grid search with a 5-fold cross validation approach has been employed which splits the training data into n sets and utilizes the created sets, except one, to calibrate the model and the remaining one to validate it. In particular, C varied between 1, 2, 4, 8, 16, 32, 64, 128, 256, 512, 1024, 2048, 4094 and 8192, while γ varied between 0.1, 0.2, 0.4, and 0.8. The couple of values resulting in the highest cross validation overall accuracy is selected and used for classifying the corresponding AOI. In particular, this is carried out with the pksvm tool, which is based on the largely employed open source C++ library libSVM (MCINERNEY & KEMPENEERS 2015:186). The output is a binary layer outlining the estimated urban and non-urban areas. Since 9 different configurations have been tested for generating the training points with 10 different random initializations for the 8 test sites and 2 selected time frames, overall 1440 classification maps have been produced. Moreover, for each AOI and time frame a majority voting classification map has been generated for each training configuration, summing up to a total of additional 144 maps.
To estimate which are the features that prove more relevant for discriminating urban and

non-urban areas among the 45 used, a dedicated analysis has been carried out a posteriori by using the sequential forward floating selection (SFFS) approach, which identifies and iteratively conditionally excludes the least significant features among the ones provided as input (PUDIL et al. 1994:1121f.). In particular, for each training set, the method has been applied by fixing the number of features to select equal to 10.

4.5 Accuracy assessment

In this study, a quantitative assessment of the final classification maps has been carried out by considering the overall classification accuracy and the Kappa coefficient. These statistics are calculated on the basis of an error matrix comparing for a certain number of samples, their true label (e.g., derived by photo interpretation) with the label automatically associated by the classifier. In the considered case, the problem is binary (urban vs. non-urban), thus the error matrix has a size of 2x2, where the columns describe the reference data and the rows describe the classification result.
The overall accuracy is computed as:

$$OA = \frac{\sum Correct\ Points}{\sum Reference\ Points} \tag{4.5}$$

with correct points representing the number of urban and non-urban samples for which there is agreement between the outcomes of the classifier and the real labels (i.e., those lying on the major diagonal of the error matrix).
The Kappa coefficient is calculated as:

$$\hat{K} = \frac{N \sum_{i=1}^{r} x_{ii} - \sum_{i=1}^{r} (x_{i+} * x_{+i})}{N^2 - \sum_{i=1}^{r} (x_{i+} * x_{+i})} \tag{4.6}$$

where r is the number of classes, x_{ii} the number of points in row i and column i, x_{i+} the sum of points in column i, x_{+i} the sum of points in row i, and N the sum of points in the error matrix (CONGALTON 1991:40). Compared to the overall accuracy, the Kappa coefficient considers all elements of the error matrix and not only the major diagonal ones. Further statistics that are calculated based on the error matrix are the user's accuracy (UA) and producer's accuracy (PA). The former is calculated based on the corresponding sum of the row and the agreement of the class and real label, while the latter by means of the sum of the column and the agreement of the class and real label.
Besides the employed measures, the adopted sampling method to collect reference points is of great importance, as both their number and location have a strong impact on the resulting error matrix. To this end, the Sampling Design Tool (NOAA's Biogeography Branch, Silver Spring, USA), an implementation for ArcGIS, was applied using the stratified random sampling approach. This procedure creates a shapefile with randomly placed points within

each class, whereby the amount of allocated reference points is determined by the user. The extent of the classes is defined by using one classification result as input.
Furthermore, a general guideline according to CONGALTON (2008:75) is to collect a minimum of 50 reference points for each class if the map size is less than 10,000 km². Instead, for larger maps it is advised to collect between 75 to 100 points per class. Since the size of all study areas is larger than 10,000 km² and two information classes are taken into account, then more points have been labeled to validate the classification results (equally split between the urban and non-urban classes, however in some cases more iterations were necessary due to missing reference data, resulting in uneven relation of validation points). Among the 8 study areas, the highest amount of validation points has been derived for the Pearl River Delta and Beijing, namely 1200. A total of 800 points were generated for Istanbul and Los Angeles and 700 points for Sao Paulo and Mexico City. The lowest number of points, i.e. 600, has been collected for Athens and Atlanta. Different reference points have been defined for the two considered time frames. Afterwards, their label has been assigned based on photo interpretation of VHR imagery of Google Earth and confusion matrices have been generated for all 90 classification maps for each study area and both investigated periods, as well as for the final classification maps obtained using the majority voting strategy (confusion matrices obtained using the caret package Version 6.0–64 in R). Moreover, also the enhanced GUF layers for the unsupervised extraction of training points have been validated. To this aim, 125 validation points have been generated for each test site by means of the stratified random sampling scheme.

5 Results

In this chapter, results of the intensive experimental analysis carried out for the selected 8 study sites are provided. First a statistical analysis of the obtained accuracies is given, while afterwards the final classification maps are presented and discussed along with the analysis about the effectiveness of the selected features.

5.1 Evaluation of classification results

5.1.1 Unsupervised DT classification

In Table 5.1 the confusion matrices obtained using the DT threshold based classifier for the period 2013–2015 are reported. As one can notice, both the overall accuracy and Kappa coefficient exhibit a heterogeneous behavior depending on the specific test site. Indeed, while for Mexico City, Sao Paulo and the Pearl River Delta they resulted in rather high values (with Kappa greater than 0.7), for the remaining sites they are quite poor (especially for Los Angeles and Atlanta where Kappa is lower than 0.4).
By analyzing the producer's accuracy of the urban class the underestimation of its extent is evident for the test sites of Beijing and the Pearl River Delta (despite the corresponding

Table 5.1: Confusion matrices for the results obtained from the unsupervised threshold based classification for the period 2013–2015.

Athens		Reference		
		non-u.	urban	UA
Prediction	non-u.	177	12	93.65
	urban	136	275	66.91
	PA	56.55	95.82	

$OA = 75.33\ \%$; $\hat{K} = 0.514$

Atlanta		Reference		
		non-u.	urban	UA
Prediction	non-u.	260	209	55.44
	urban	4	127	96.95
	PA	98.48	37.80	

$OA = 64.50\ \%$; $\hat{K} = 0.335$

Beijing		Reference		
		non-u.	urban	UA
Prediction	non-u.	558	213	72.37
	urban	32	397	92.54
	PA	94.58	65.08	

$OA = 79.58\ \%$; $\hat{K} = 0.594$

Istanbul		Reference		
		non-u.	urban	UA
Prediction	non-u.	345	55	86.25
	urban	72	328	82.00
	PA	82.73	85.64	

$OA = 84.13\ \%$; $\hat{K} = 0.683$

Los Angeles		Reference		
		non-u.	urban	UA
Prediction	non-u.	147	7	95.45
	urban	234	412	63.78
	PA	38.58	98.33	

$OA = 69.88\ \%$; $\hat{K} = 0.379$

Mexico City		Reference		
		non-u.	urban	UA
Prediction	non-u.	304	52	85.39
	urban	47	297	86.34
	PA	86.61	85.10	

$OA = 85.86\ \%$; $\hat{K} = 0.717$

PRD		Reference		
		non-u.	urban	UA
Prediction	non-u.	601	99	85.86
	urban	55	445	89.00
	PA	91.62	81.80	

$OA = 87.17\ \%$; $\hat{K} = 0.739$

Sao Paulo		Reference		
		non-u.	urban	UA
Prediction	non-u.	333	55	85.82
	urban	13	299	95.83
	PA	96.24	84.46	

$OA = 90.29\ \%$; $\hat{K} = 0.806$

overall accuracy is about 80 % in both cases), as well as for Atlanta (where just 127 validation points over 336 are correctly marked as urban). On the contrary, an overestimation of urban areas occurs for Athens, Istanbul, and Los Angeles (where just 234 of the available 381 validation points of the non-urban class are classified correctly, resulting in a producer's accuracy equal to 38.58 %). The accuracy measures and the classification results clearly indicate that the classification problem is to complex to solve it with a simple threshold based linear classifier.

5.1.2 Proposed classification scheme

Contrarily to the simple DT classifier, the proposed SVM based system proved extremely robust and resulted in very high accuracies for all 8 test sites and both investigated time frames. First, to prove the reliability of the implemented strategy for automatically extracting labeled points from the GUF, the accuracies of the following 'training' masks generated for each AOI have been computed and reported in Table 5.2. In particular, according with the description provided in Section 4.4.2.1:

- for the 2013–2015 period a mask has been derived where for the urban class the only GUF version 1 pixels remaining after the filtering are kept, while for the non-urban class the only pixels outside the buffer created along the fringes of the GUF version 7 urban areas are preserved (all the remaining areas are then excluded from the analysis);

- for the 2002–2003 period a mask has been derived where for the urban class the only GUF version 1 pixels satisfying the criteria on the temporal NDBI and NDVI standard deviation are kept, while for the non-urban class the only pixels outside the (smaller) buffer created along the fringes of the GUF version 7 urban areas are preserved (all the remaining areas are then excluded from the analysis).

Table 5.2: Accuracy of the generated 'training' masks for automatic collection of training points.

Athens	OA	$\hat{K}$
Mask 2002–2003	98.80	0.976
Mask 2013–2015	99.60	0.992
Beijing	OA	$\hat{K}$
Mask 2002–2003	97.20	0.944
Mask 2013–2015	99.20	0.984
Los Angeles	OA	$\hat{K}$
Mask 2002–2003	98.40	0.968
Mask 2013–2015	99.20	0.984
Pearl River Delta	OA	$\hat{K}$
Mask 2002–2003	98.00	0.960
Mask 2013–2015	99.60	0.992

Atlanta	OA	$\hat{K}$
Mask 2002–2003	96.80	0.936
Mask 2013–2015	98.40	0.968
Istanbul	OA	$\hat{K}$
Mask 2002–2003	97.20	0.944
Mask 2013–2015	99.60	0.992
Mexico City	OA	$\hat{K}$
Mask 2002–2003	98.80	0.976
Mask 2013–2015	99.20	0.984
Sao Paulo	OA	$\hat{K}$
Mask 2002–2003	98.00	0.960
Mask 2013–2015	99.60	0.992

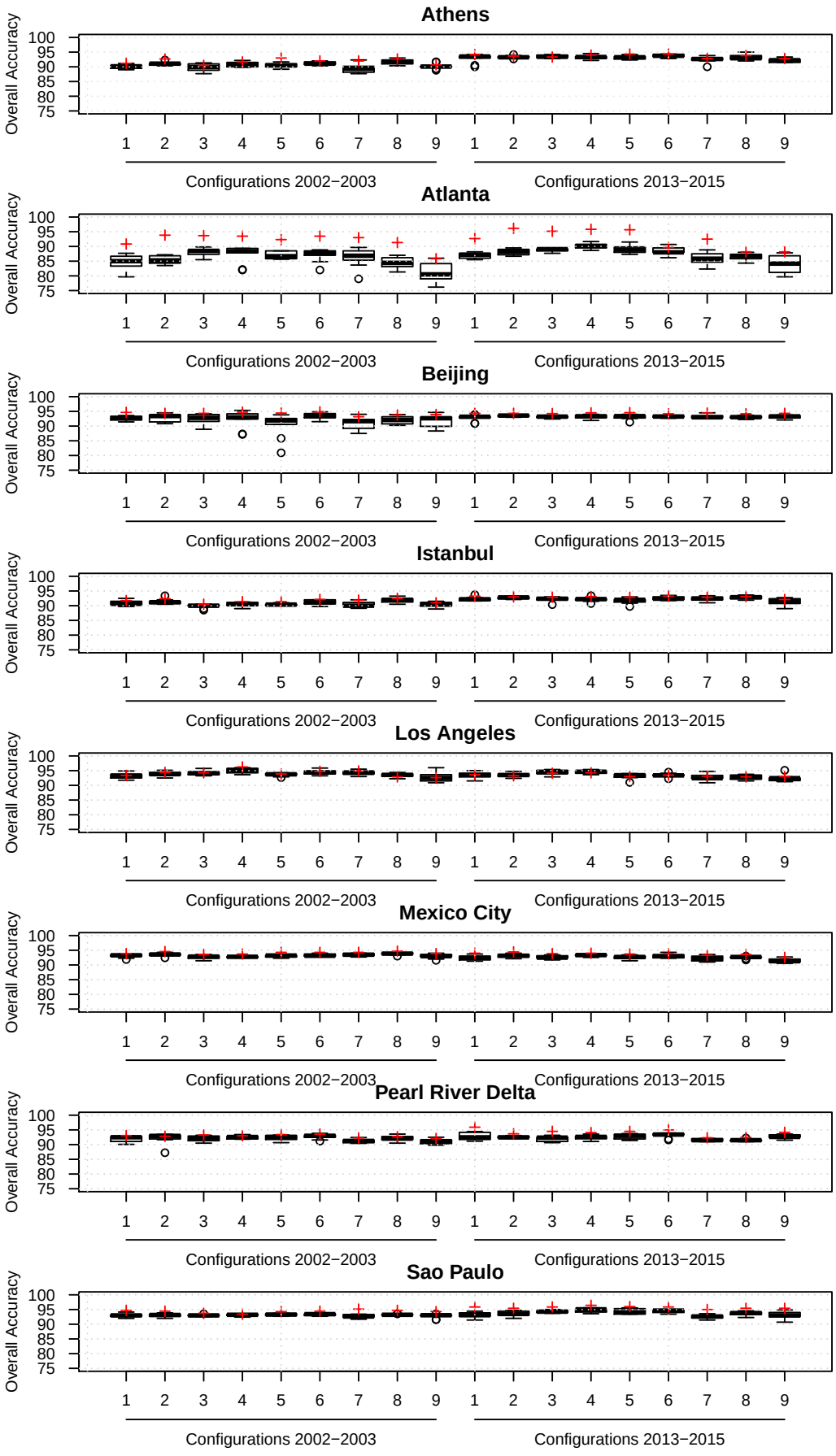

Figure 5.1: Boxplots showing the resulting overall accuracy for each classification configuration, while the red crosses represent the accuracy statistics for the final classifications obtained from the majority voting strategy.

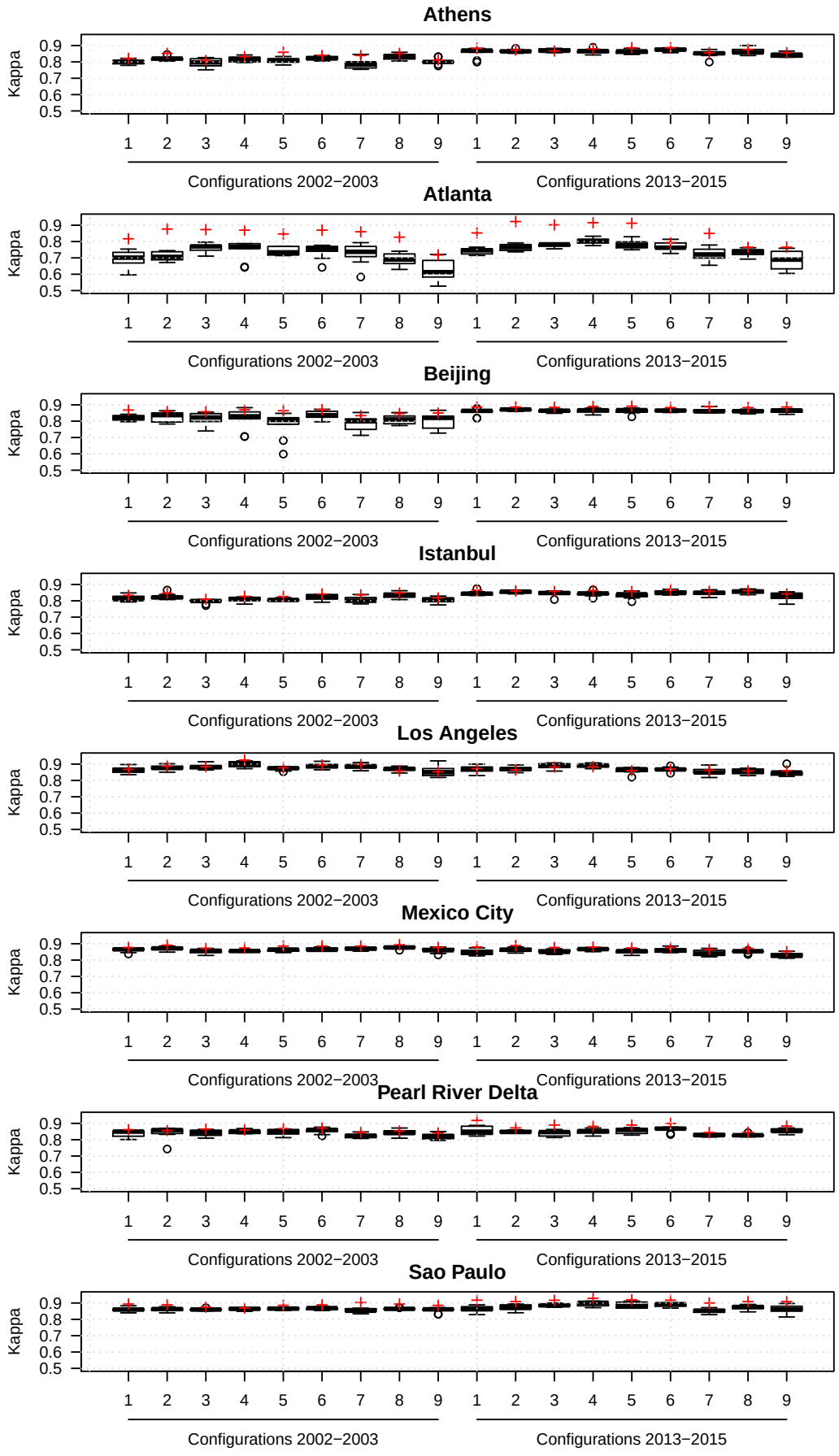

Figure 5.2: Boxplots showing the resulting Kappa values for each classification configuration, while the red crosses represent the accuracy statistics for the final classifications obtained from the majority voting strategy.

Table 5.3: Confusion matrices for the results obtained from the SVM based classification for the period 2002–2003.

Athens

Prediction		Reference non-u.	Reference urban	UA
	non-u.	292	21	93.29
	urban	29	258	89.90
	PA	90.97	92.47	

$OA = 91.67\ \%$; $\hat{K} = 0.833$

Atlanta

Prediction		Reference non-u.	Reference urban	UA
	non-u.	269	16	94.39
	urban	23	292	92.70
	PA	92.12	94.81	

$OA = 93.50\ \%$; $\hat{K} = 0.870$

Beijing

Prediction		Reference non-u.	Reference urban	UA
	non-u.	306	25	92.45
	urban	38	831	95.63
	PA	88.95	97.08	

$OA = 94.75\ \%$; $\hat{K} = 0.870$

Istanbul

Prediction		Reference non-u.	Reference urban	UA
	non-u.	401	5	98.77
	urban	64	330	83.76
	PA	86.24	98.51	

$OA = 91.38\ \%$; $\hat{K} = 0.827$

Los Angeles

Prediction		Reference non-u.	Reference urban	UA
	non-u.	359	18	95.23
	urban	12	411	97.16
	PA	96.77	95.80	

$OA = 96.25\ \%$; $\hat{K} = 0.925$

Mexico City

Prediction		Reference non-u.	Reference urban	UA
	non-u.	315	11	96.63
	urban	33	341	91.18
	PA	90.52	96.88	

$OA = 93.71\ \%$; $\hat{K} = 0.874$

PRD

Prediction		Reference non-u.	Reference urban	UA
	non-u.	574	40	93.49
	urban	44	542	92.49
	PA	92.88	93.13	

$OA = 93.00\ \%$; $\hat{K} = 0.860$

Sao Paulo

Prediction		Reference non-u.	Reference urban	UA
	non-u.	327	9	97.32
	urban	37	327	89.84
	PA	89.84	97.32	

$OA = 93.43\ \%$; $\hat{K} = 0.869$

For the masks of the time frame 2013–2015, the overall accuracy and the Kappa coefficient are always higher than 99.20 % and 0.968, respectively, whereas for the masks of 2002–2003 they proved to be always greater than 96.80 % and 0.936, respectively. In general, the corresponding randomly generated training sets are then expected to be correct, with the exception – in case – of some few pixels. However, this proved not critical being the SVM capable of automatically limiting their influence in determining the final classification function by means of the penalization parameter C.

The boxplots in Figure 5.1 and Figure 5.2 depict the statistical distribution of the overall accuracy and Kappa coefficient, respectively, obtained from the 10 randomly generated training sets for each of the considered 9 sampling configurations. For each site, the left and right parts refer to the 2002–2003 and 2013–2015 periods, respectively. Moreover, the results derived applying the majority voting strategy are also reported as red crosses.

The reduced spread of all boxes (apart from very few exceptions) highlights the robustness versus different training initializations, which then have a limited effect on the final results. Only for the AOIs of Atlanta and Beijing the variability among the 9 sampling configurations is consistent. In the former case, this occurs for both 2002–2003 and 2013–2015 and it is mostly due to the fact that Atlanta is characterized by many sub-urban low-density residential areas with buildings often hidden by vegetation; hence, configurations with highly unbalanced number of training samples tend to result in underestimation of the

urban class. In the latter case, the phenomenon solely occurs in 2002–2003. However, it is worth pointing out that between 2002–2003 and 2011–2013 (the period which the GUF refers to) Beijing experienced an unprecedented growth. Accordingly, it might occur that, despite the pruning based on the temporal NDBI and NDVI standard deviation to derive the 'training' mask for 2002–2003, some training samples are wrongly marked as urban which, also here, might lead to lower performances in sampling configurations with highly unbalanced number of labeled points for the two classes. For this same reason, accuracies are slightly higher for the period 2013–2015 with respect to 2002–2003 also for Istanbul. This holds also for Athens, but it is actually due to the availability of a lower number of scenes (i.e., 59 in 2002–2003 vs. 96 in 2013–2015).

In general, the configurations with 1,000 training points for both urban and non-urban classes allowed to obtain the best performances with an average overall accuracy and Kappa over the whole test sites equal to 93.46 % and 0.866 for 2002–2003 and 94.53 % and 0.891 for 2013–2015, respectively.

Using the majority voting strategy proved always very effective. This is evident when analyzing the case of Atlanta, where it led to very good results (despite single classifications were not highly accurate) with an improvement of overall accuracy and Kappa greater than 10 % and 0.1, respectively. Moreover, this also confirms that, as assumed, errors obtained with different training initializations tend to be uncorrelated.

Table 5.4: Confusion matrices for the results obtained from the SVM based classification for the period 2013–2015.

Athens		Reference		
		non-u.	urban	UA
Prediction	non-u.	286	9	96.95
	urban	27	278	91.15
	PA	91.37	96.86	

$OA = 94.00\ \%$; $\hat{K} = 0.880$

Atlanta		Reference		
		non-u.	urban	UA
Prediction	non-u.	251	12	95.44
	urban	13	324	96.14
	PA	95.08	96.43	

$OA = 95.83\ \%$; $\hat{K} = 0.915$

Beijing		Reference		
		non-u.	urban	UA
Prediction	non-u.	543	19	96.62
	urban	47	591	92.63
	PA	92.03	96.89	

$OA = 94.50\ \%$; $\hat{K} = 0.890$

Istanbul		Reference		
		non-u.	urban	UA
Prediction	non-u.	372	10	97.38
	urban	45	373	89.23
	PA	89.21	97.39	

$OA = 93.13\ \%$; $\hat{K} = 0.869$

Los Angeles		Reference		
		non-u.	urban	UA
Prediction	non-u.	352	17	95.39
	urban	29	402	93.27
	PA	92.39	95.94	

$OA = 94.25\ \%$; $\hat{K} = 0.885$

Mexico City		Reference		
		non-u.	urban	UA
Prediction	non-u.	322	13	96.12
	urban	29	336	92.05
	PA	91.74	96.28	

$OA = 94.00\ \%$; $\hat{K} = 0.880$

PRD		Reference		
		non-u.	urban	UA
Prediction	non-u.	592	7	98.83
	urban	64	537	89.35
	PA	90.24	98.71	

$OA = 94.08\ \%$; $\hat{K} = 0.882$

Sao Paulo		Reference		
		non-u.	urban	UA
Prediction	non-u.	330	9	97.35
	urban	16	345	95.57
	PA	95.38	97.46	

$OA = 96.43\ \%$; $\hat{K} = 0.929$

Table 5.3 and Table 5.4 report the confusion matrices derived applying the majority voting to the classification maps obtained with the best sampling configuration for 2002–2003 and 2013–2015, respectively. For the period 2002–2003 the overall accuracy is always higher than 91 % and the Kappa greater than 0.82 with a peak for Los Angeles (overall accuracy equal to 96.25 % and Kappa equal to 0.925). As mentioned above, results for the period 2013–2015 are slightly higher (also due to the fact that the temporal gap with respect to the GUF is shorter). The highest accuracies have been obtained for Sao Paulo (overall accuracy equal to 96.43 % and Kappa equal to 0.929), and, in general, the overall accuracy is always higher than 93 % and the Kappa greater than 0.86. All these numbers assess the great capabilities of the implemented system and its effectiveness in outlining urban areas, as also shown in the next Section where the final classification maps are given.

5.2 Urban extent maps

Figure 5.3 to Figure 5.6 depict the classification maps obtained using the threshold based DT classification approach. As discussed in the previous Section, the corresponding accuracy assessment proved that their quality consistently varies across the addressed study areas. Simply by visually comparing the estimated extent with VHR imagery, the overestimation of urban areas is clearly visible for the study areas of Athens and Los Angeles where mountainous areas and barren land are mislabeled as urban. At the same time, it is also easy to notice the underestimation for Atlanta and Beijing mostly occurring in low-density urban areas and throughout the entire study area especially in dense urban areas, respectively. Instead, the maps for Mexico City, the Pearl River Delta, and Sao Paulo well outline the current urban extent.

The maps derived with the proposed SVM based system are given as of Figure 5.7. Specifically, they refer to the case where the majority voting approach has been applied to the single classifications obtained using the sampling configuration 4. In particular, the 2002–2003 extent is depicted in foreground with a lighter tone, whereas that for 2013–2015 is shown with a darker tone in background. This allows to clearly identify the areas experiencing a growth between the two considered periods. By comparing the results with the corresponding available VHR Google Earth imagery, it is possible to appreciate their very high accuracy. Only dry riverbeds, mineral extraction sites, and bare rocks tend (solely under specific conditions) to be misclassified as urban areas. However, this occurs seldom and is due to the fact that, in such cases, the corresponding spectral signature in the Landsat bands is extremely similar to that of the urban class.

Athens experienced only small changes between the two analyzed time frames, whereas both in Los Angeles and Atlanta a consistent growth occurred but mostly in the outer metropolitan area (while the urban extent within the city boundaries basically remained unchanged). As already mentioned, the urban expansion in Beijing is remarkable; in par-

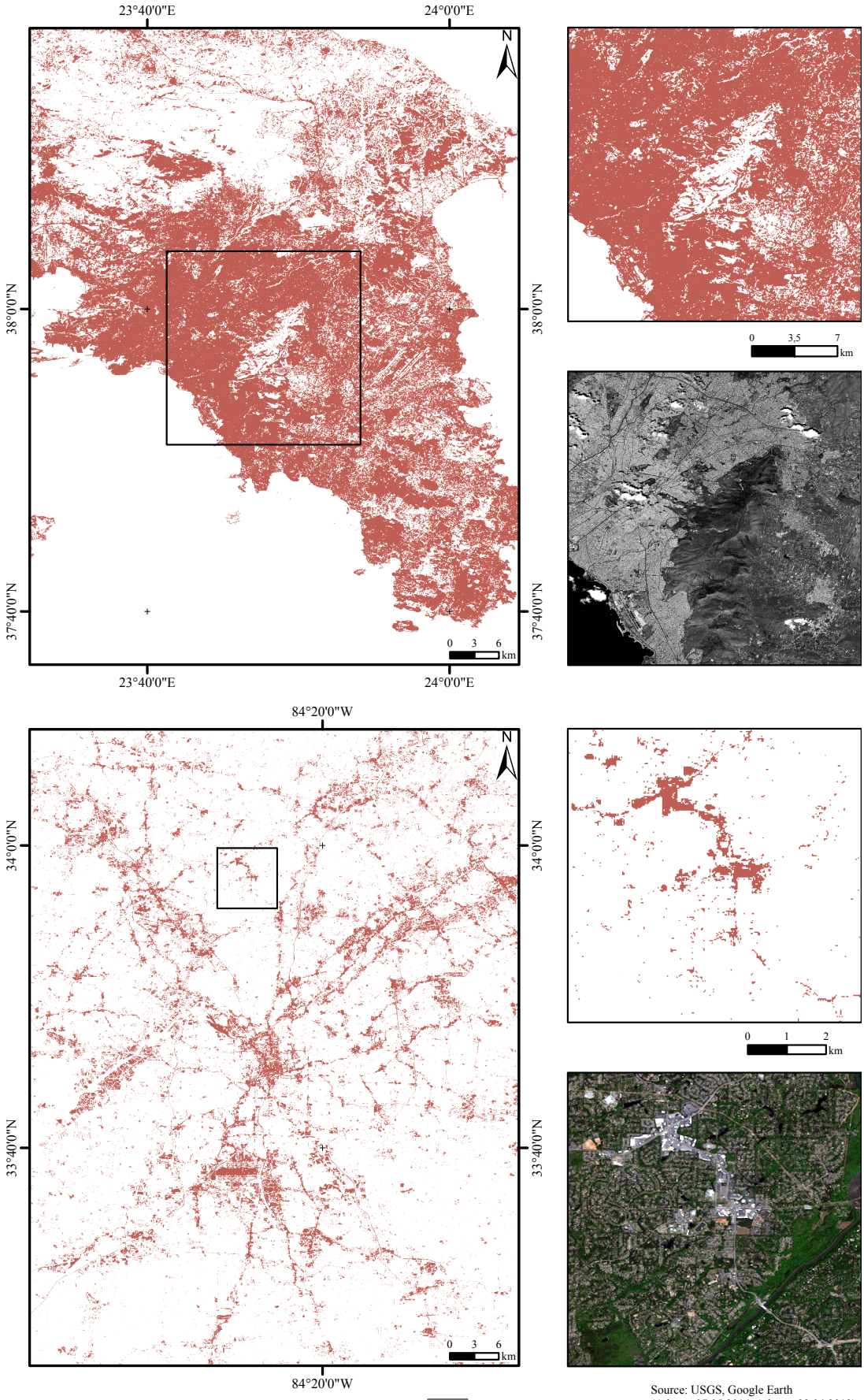

Figure 5.3: Urban extent maps for Athens and Atlanta based on DT classification scheme.

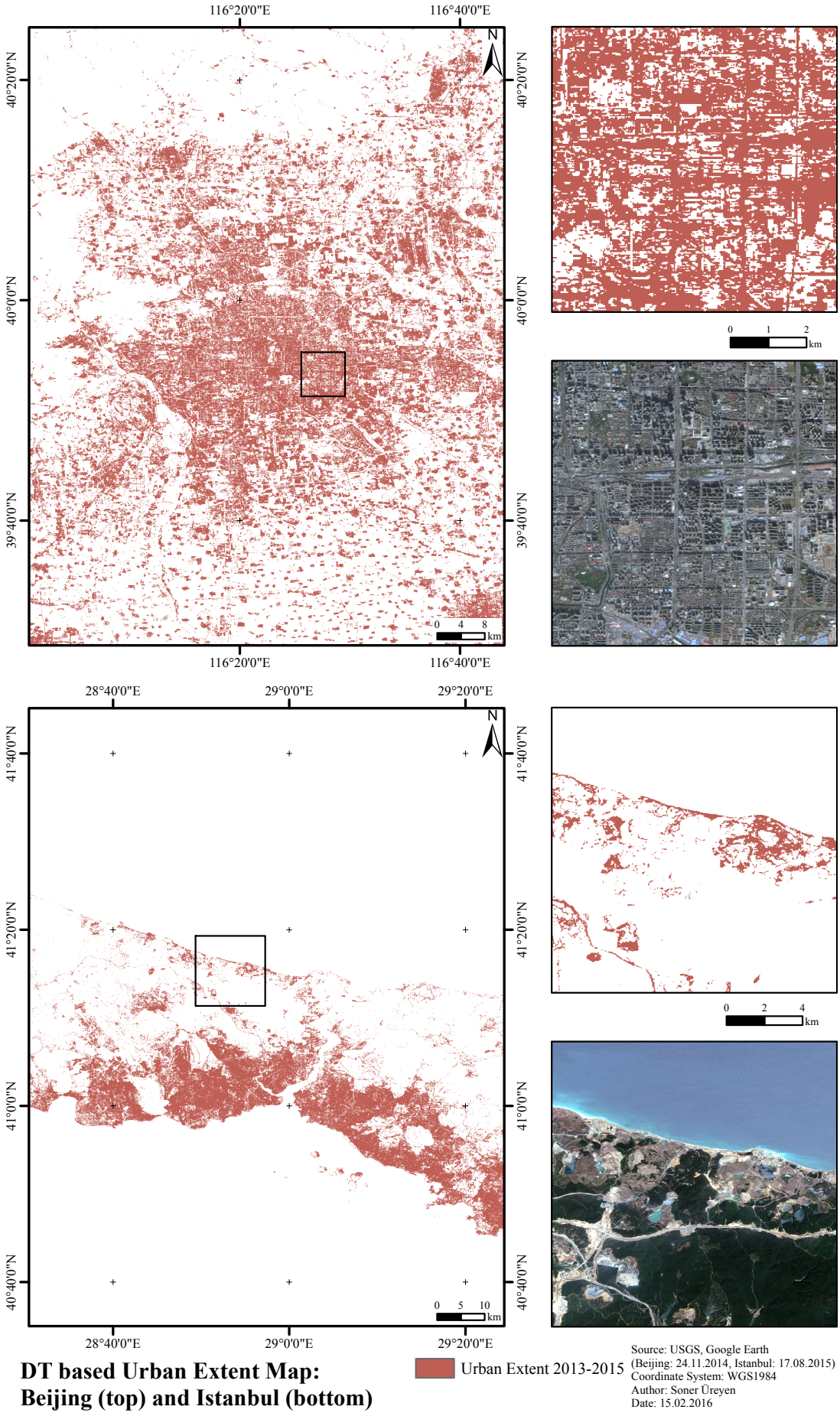

Figure 5.4: Urban extent maps for Beijing and Istanbul based on DT classification scheme.

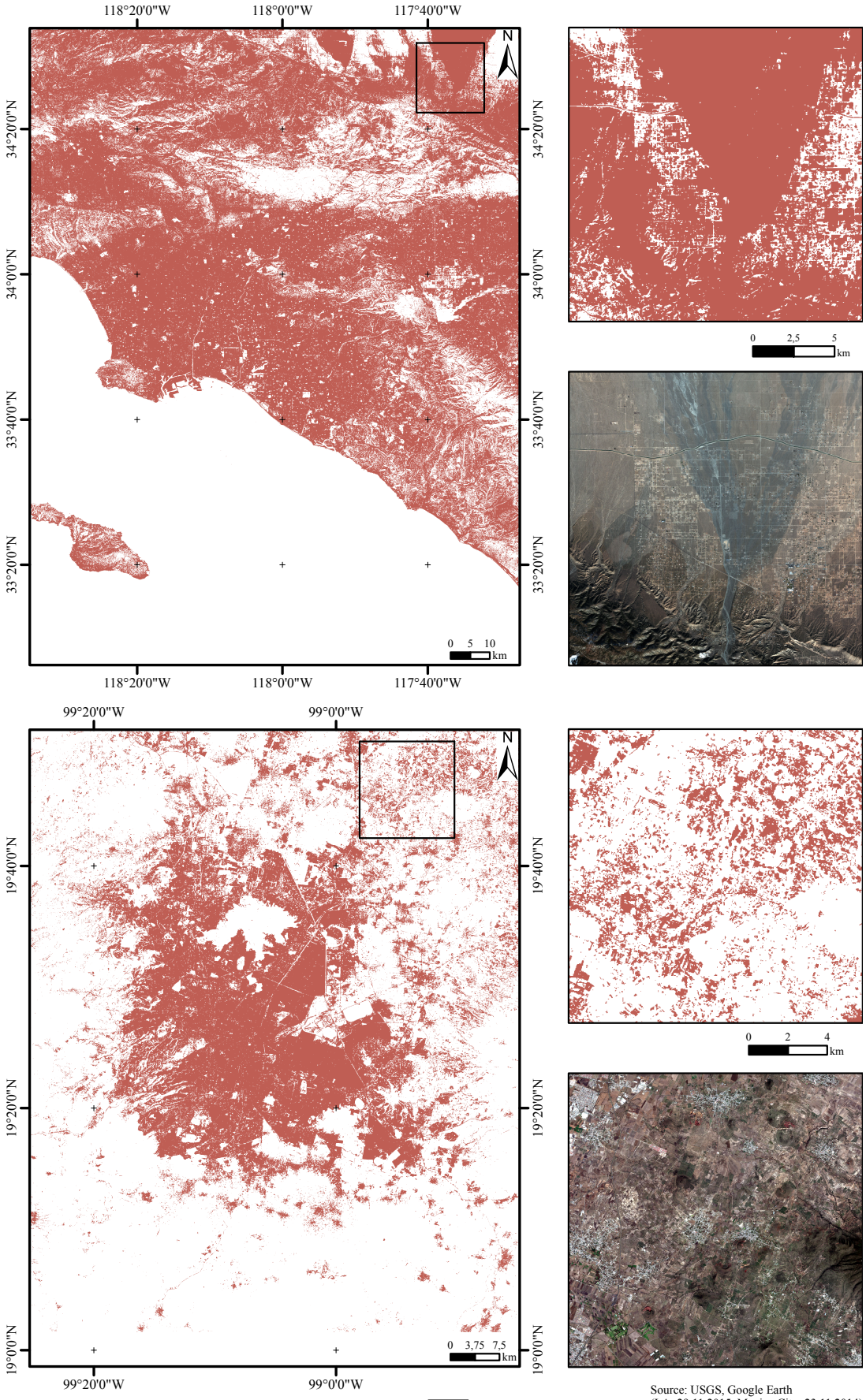

Figure 5.5: Urban extent maps for Los Angeles and Mexico City based on DT classification scheme.

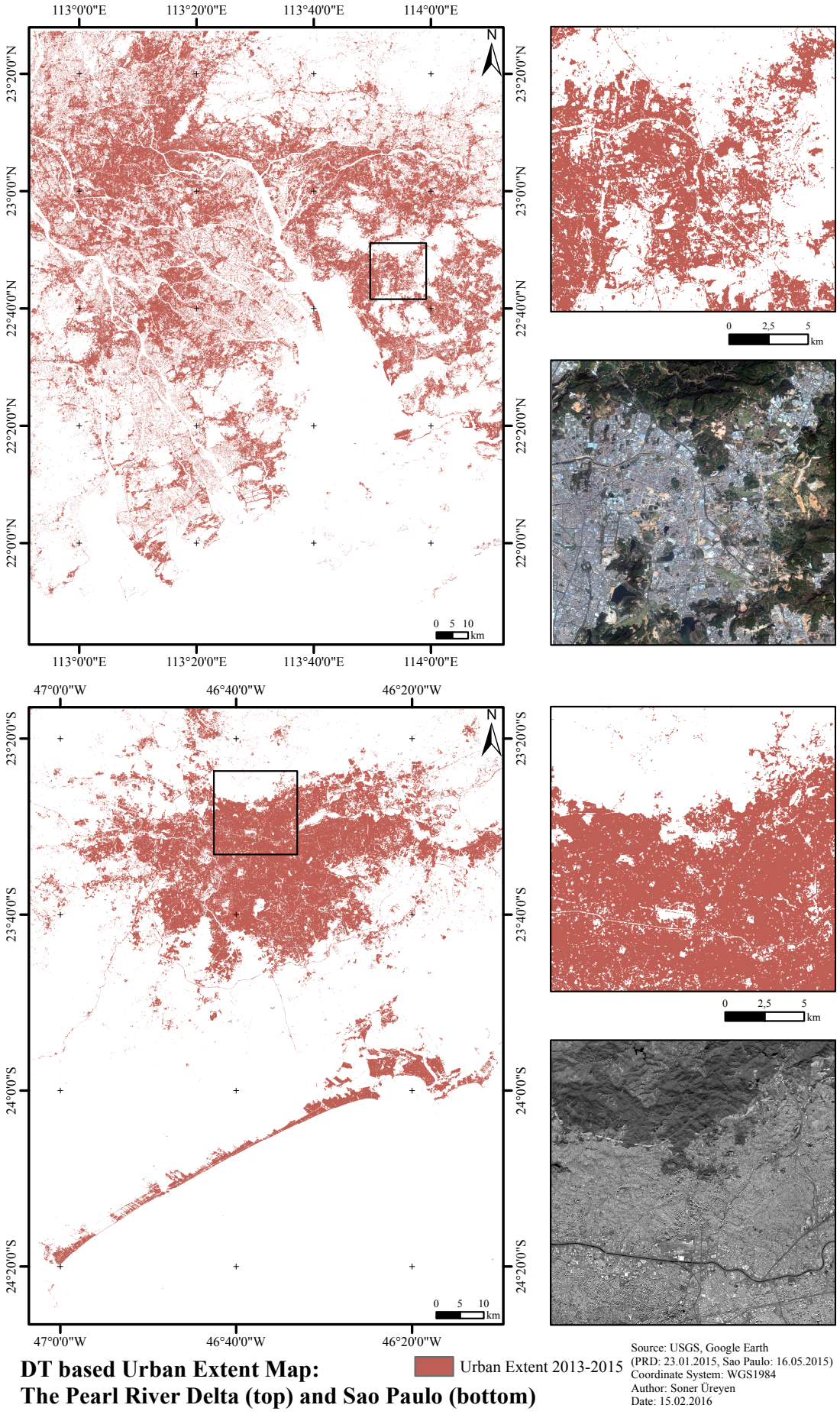

Figure 5.6: Urban extent maps for the Pearl River Delta and Sao Paulo based on DT classification scheme.

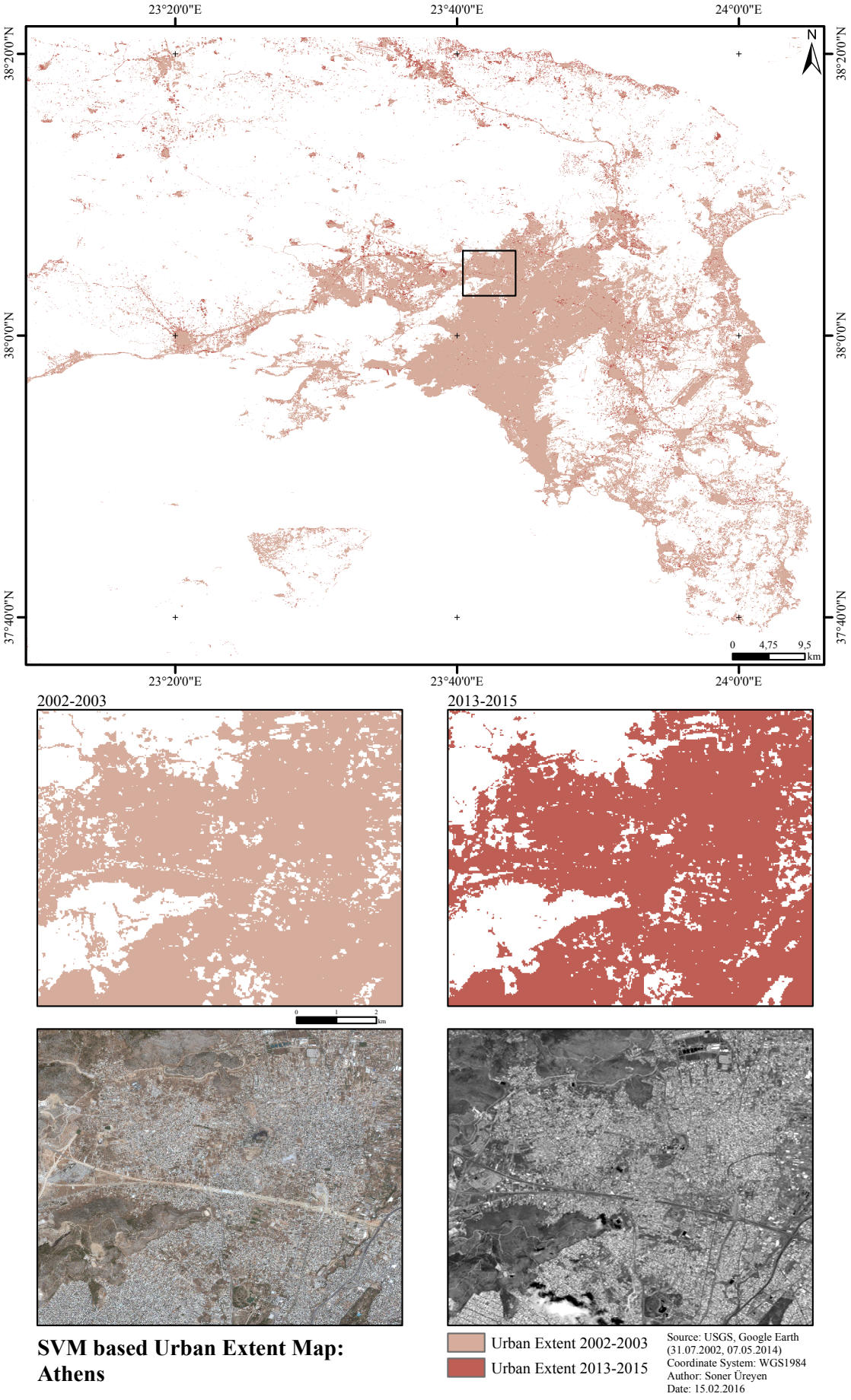

Figure 5.7: Urban extent maps for Athens based on proposed SVM classification scheme.

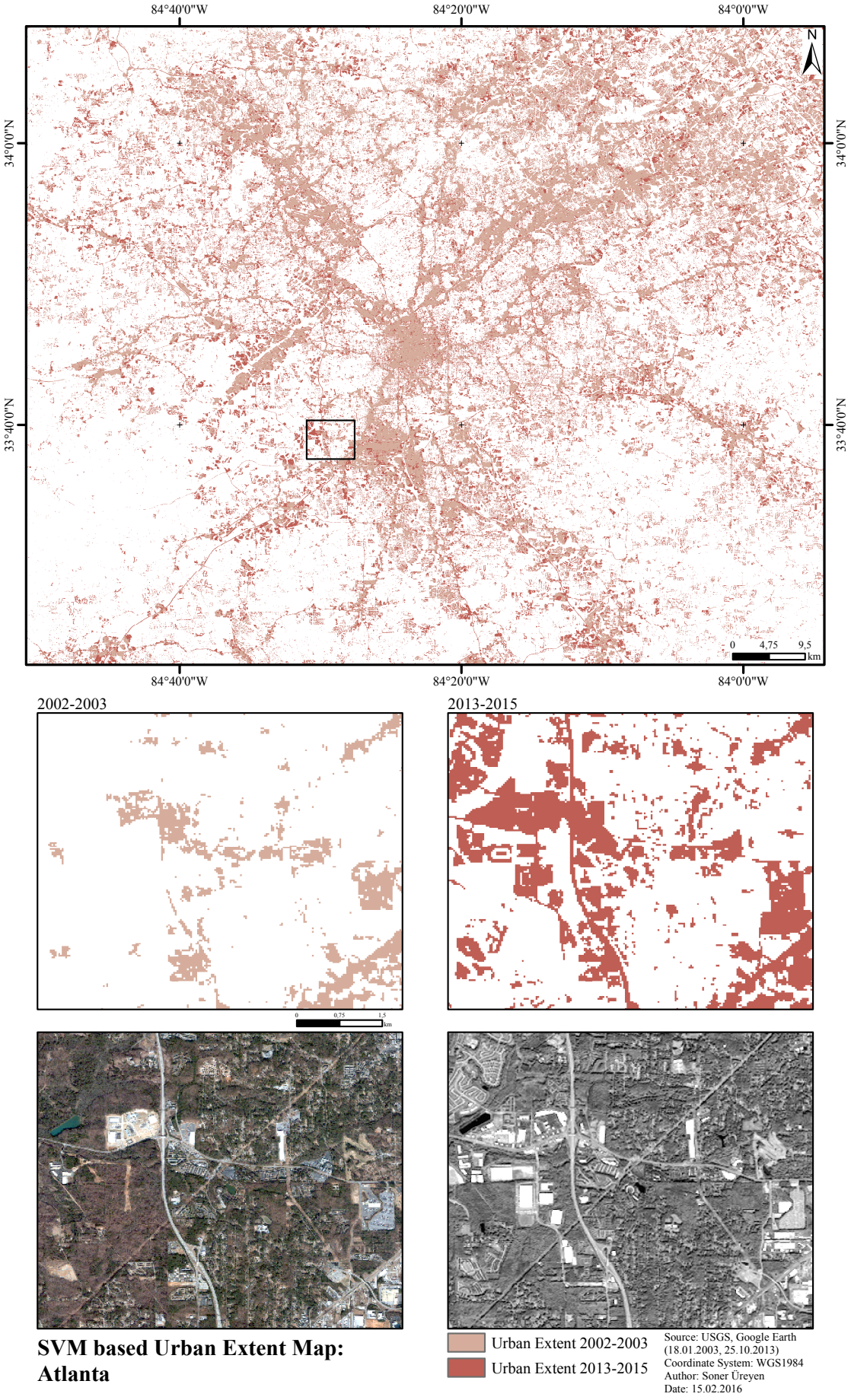

Figure 5.8: Urban extent maps for Atlanta based on proposed SVM classification scheme.

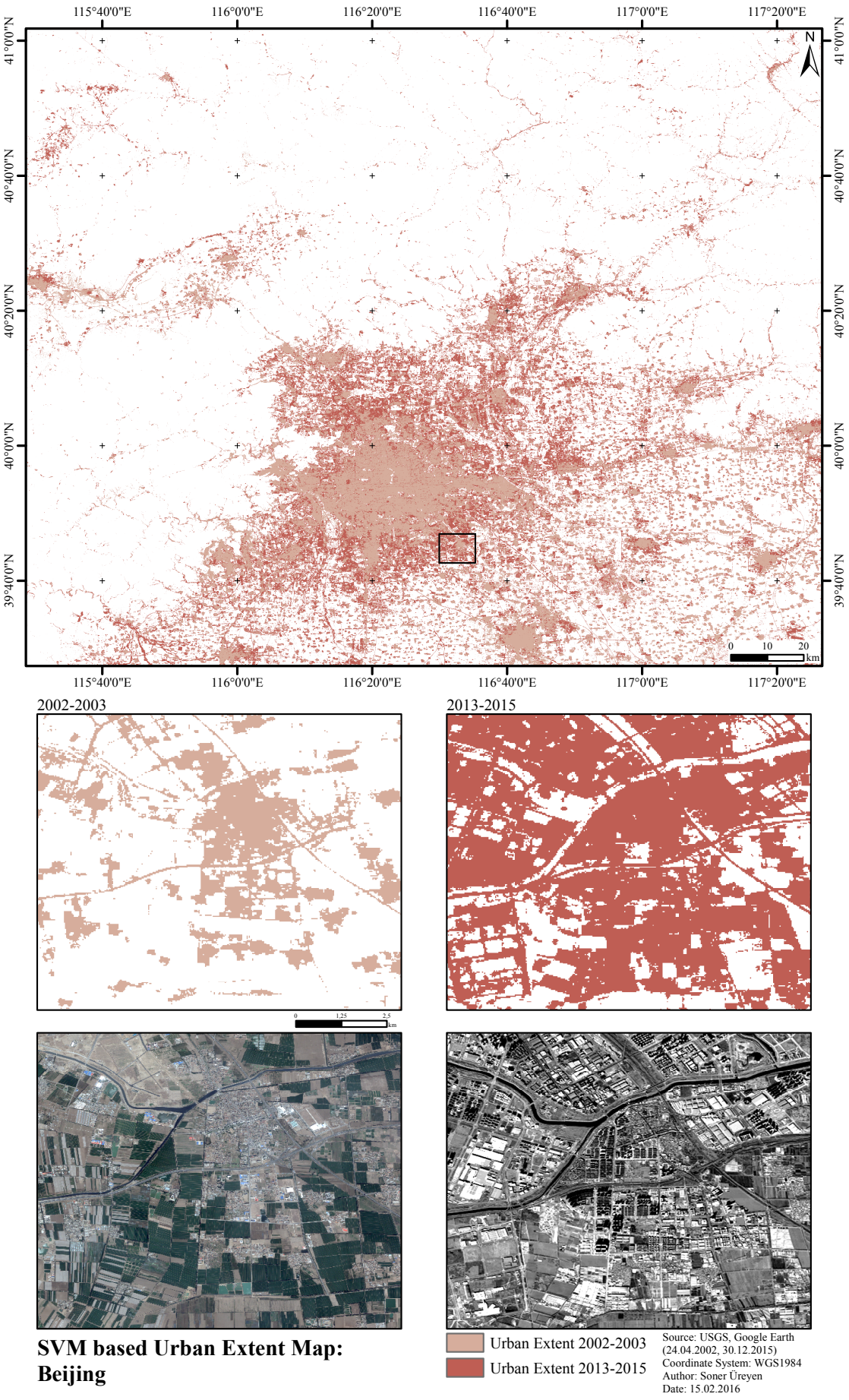

Figure 5.9: Urban extent maps for Beijing based on proposed SVM classification scheme.

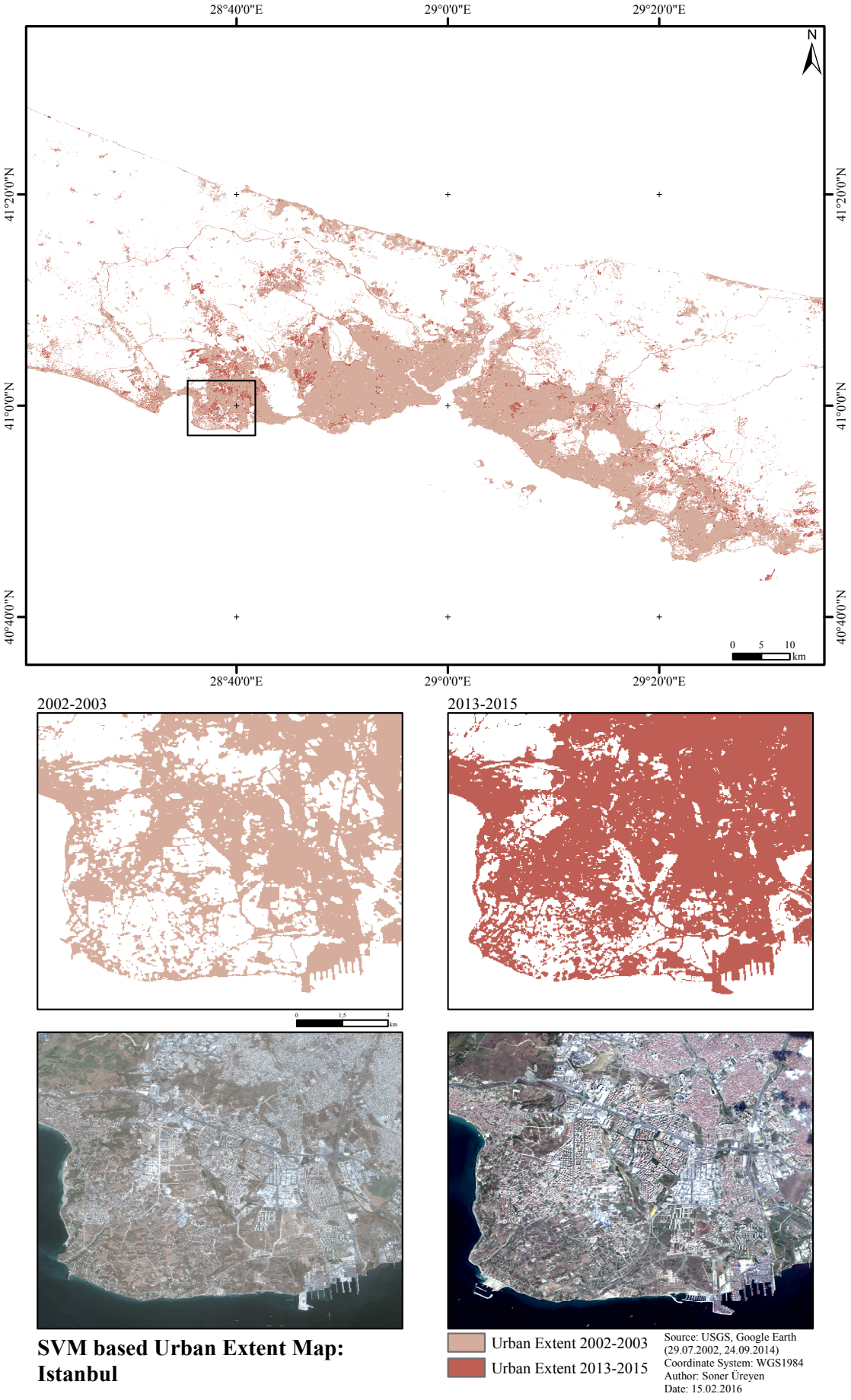

Figure 5.10: Urban extent maps for Istanbul based on proposed SVM classification scheme.

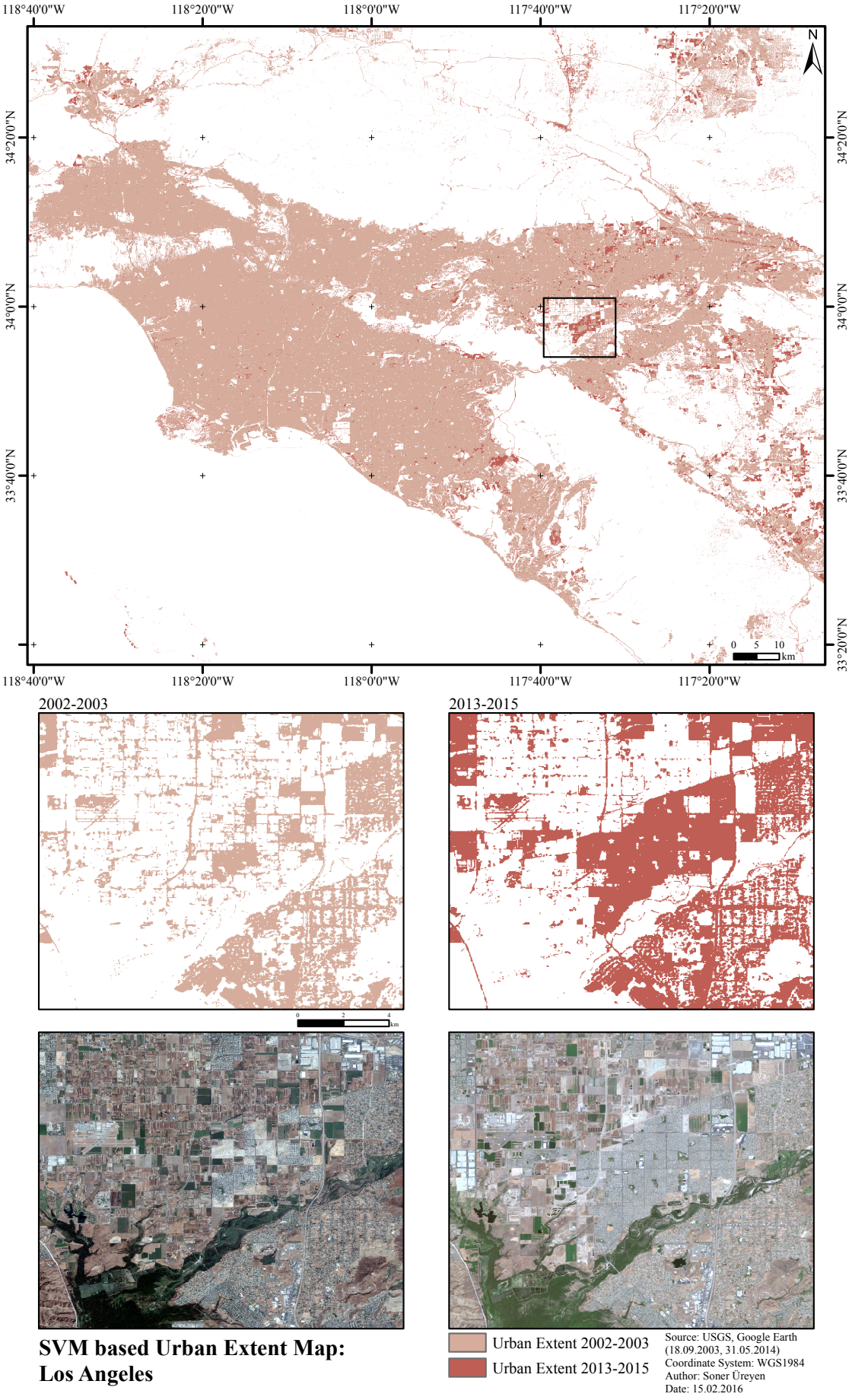

Figure 5.11: Urban extent maps for Los Angeles based on proposed SVM classification scheme.

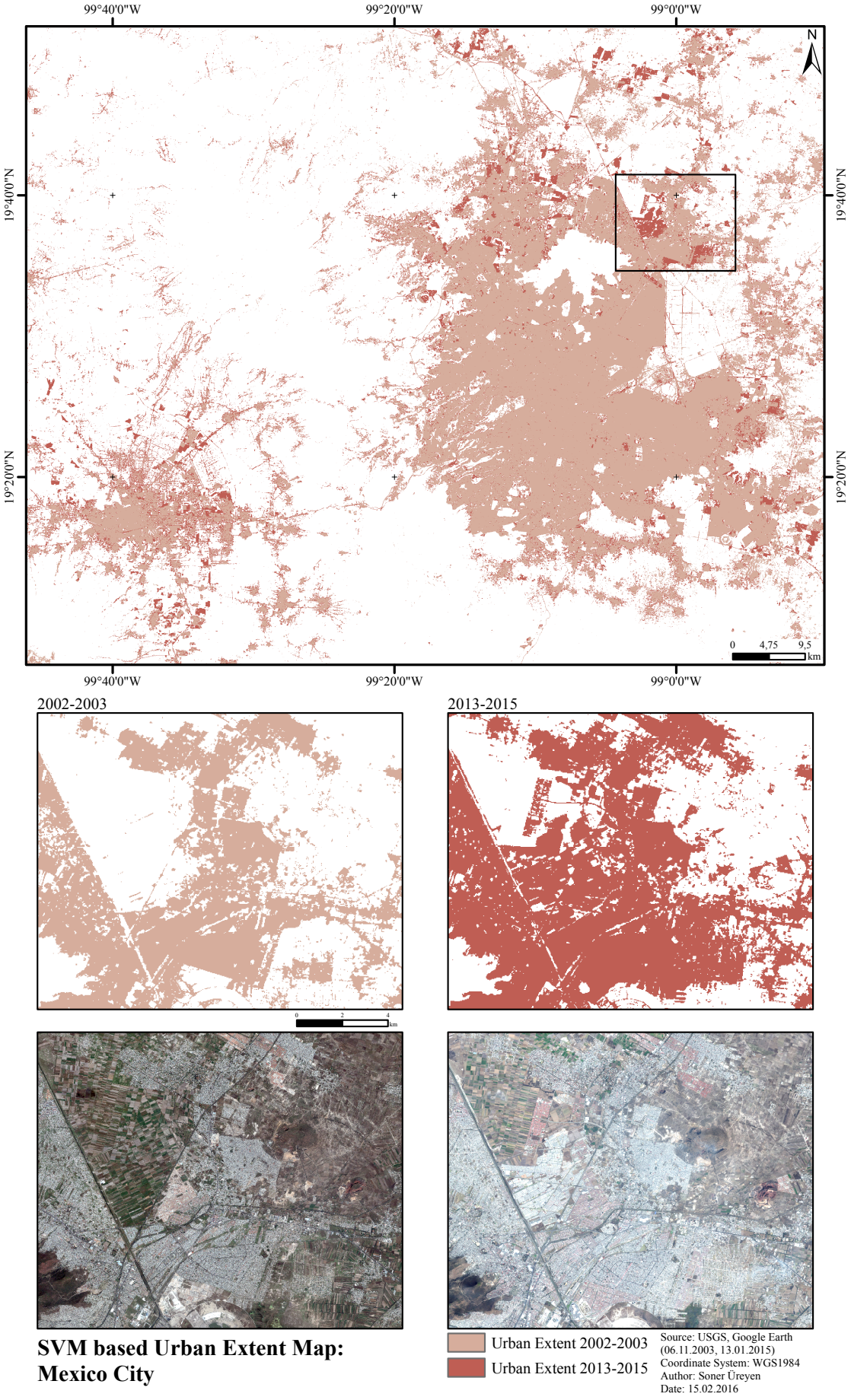

Figure 5.12: Urban extent maps for Mexico City based on proposed SVM classification scheme.

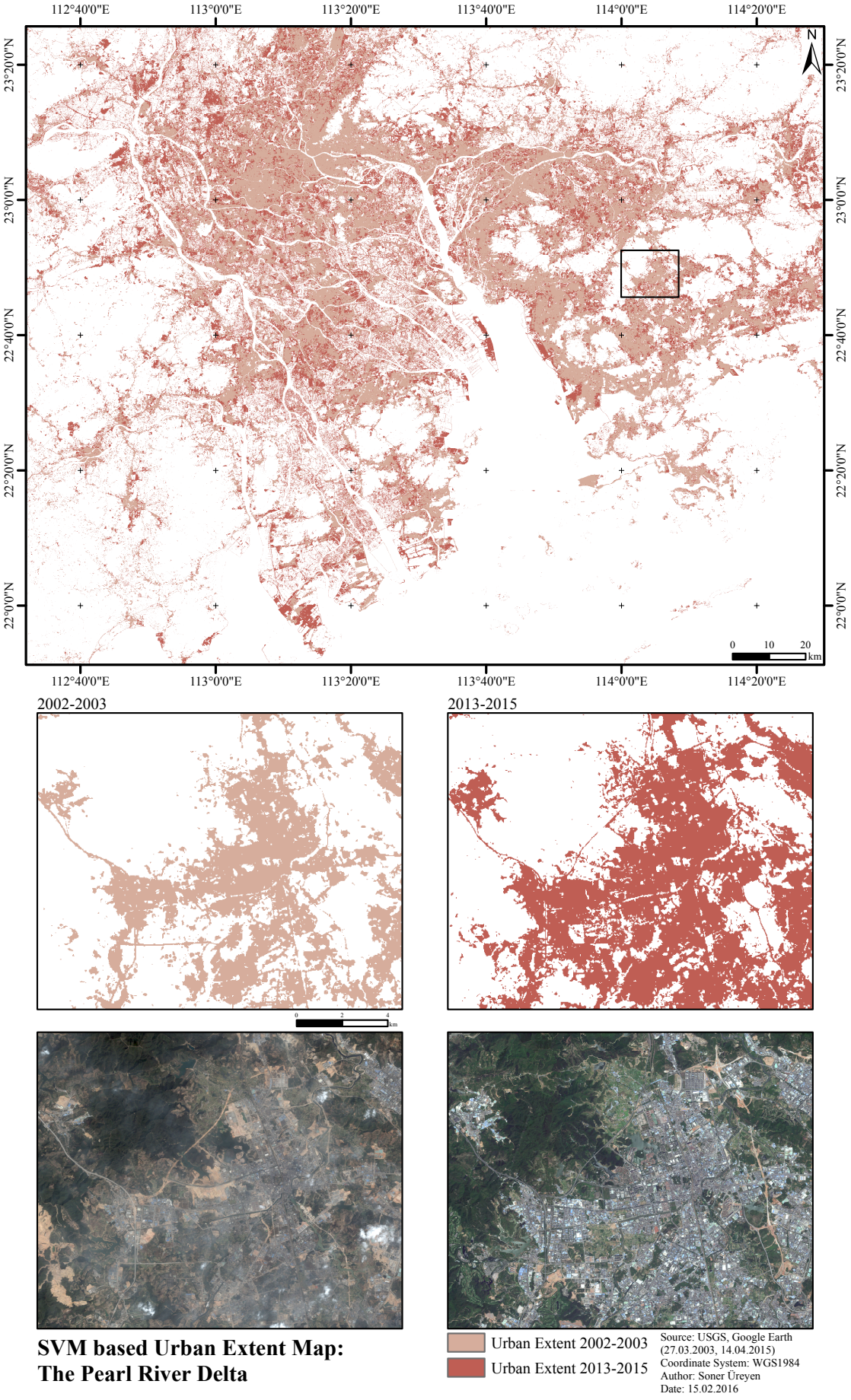

Figure 5.13: Urban extent maps for the Pearl River Delta based on proposed SVM classification scheme.

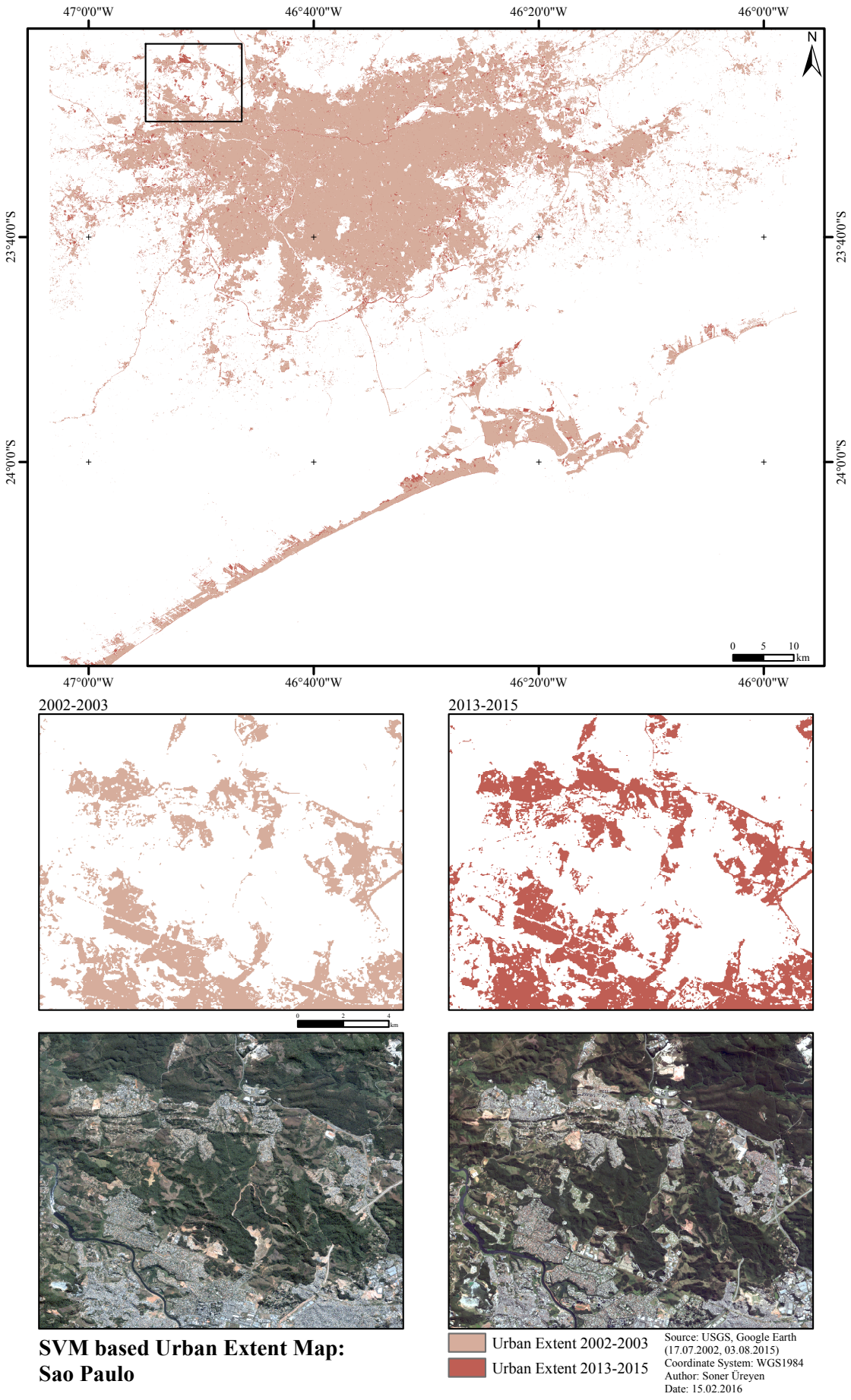

Figure 5.14: Urban extent maps for Sao Paulo based on proposed SVM classification scheme.

ticular, due to the presence of mountains at the northern and western side of the city, the growth took place eastwards and, above all, southwards. In the case of Istanbul, the growth mostly occurred on the European side of the city, while in Mexico City new urban areas have been built predominantly in the Northern side of the investigated AOI. The urban area in the Pearl River Delta sensibly increased from 2002–2003, leading to the appearance of the current biggest megaregion on the Earth. Finally, despite its big size, only few changes occurred in the Sao Paulo site.

5.3 Monitoring of urban growth

The bar graph in Figure 5.15 reports the estimated extent of the urban areas in km² for each study site and both considered time periods. It is worth noting that the values are not related to the size of the administrative boundaries of the cities but rather to the entire investigated AOI. Accordingly, the Pearl River Delta experienced the most significant growth, with the total surface of urban areas more than doubled between 2002–2003 and 2013–2015. A similar trend, despite at a lower scale, occurred in Beijing. Among the remaining cities, Atlanta exhibited a relatively consistent urban extent increase, while the change in Athens and Sao Paulo is rather limited.

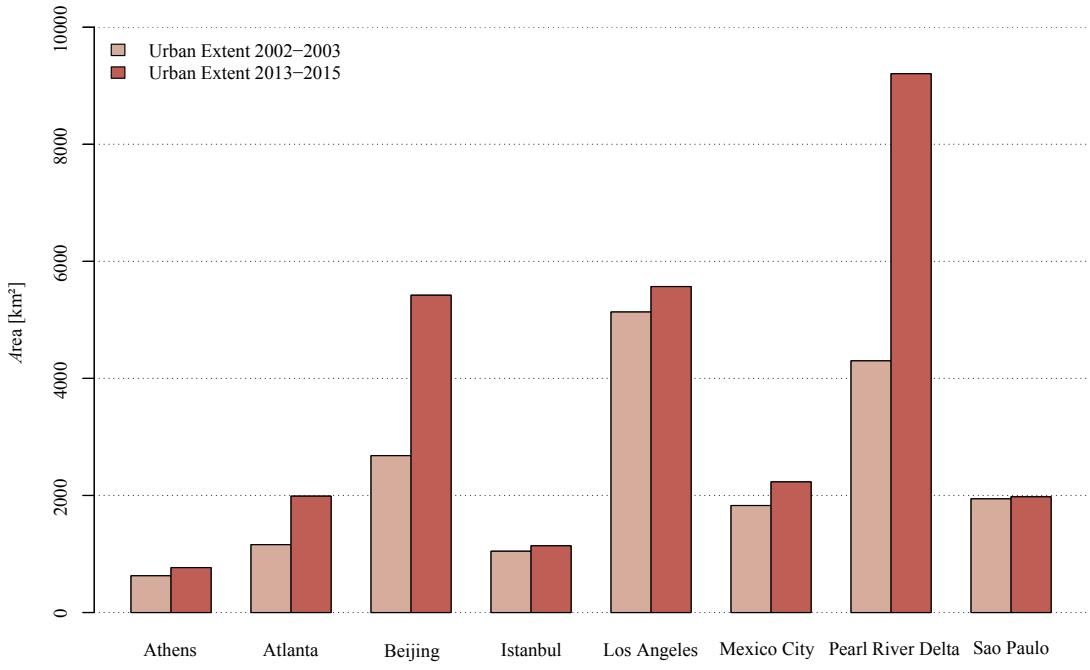

Figure 5.15: Changes in urban area extent for the 8 study areas between 2002–2003 and 2013–2015 obtained from the SVM based classification scheme.

5.4 Feature importance

An analysis for assessing the most relevant features among the 45 extracted for addressing the considered classification problem has been carried out for each test city and period of interest using the SFFS algorithm. In particular, the SFFS is run over the 10 randomly extracted training sets for each of the 9 tested sampling configurations, for a total of 90 times. The features are then sorted based on the number of times where they have been selected among the 10 most effective for each training set. In Figure 5.16 the overall ranking list of the first 25 features is given for each study area and time frame.
In most of the cases, the temporal mean indices are ranked higher than the other statistics,

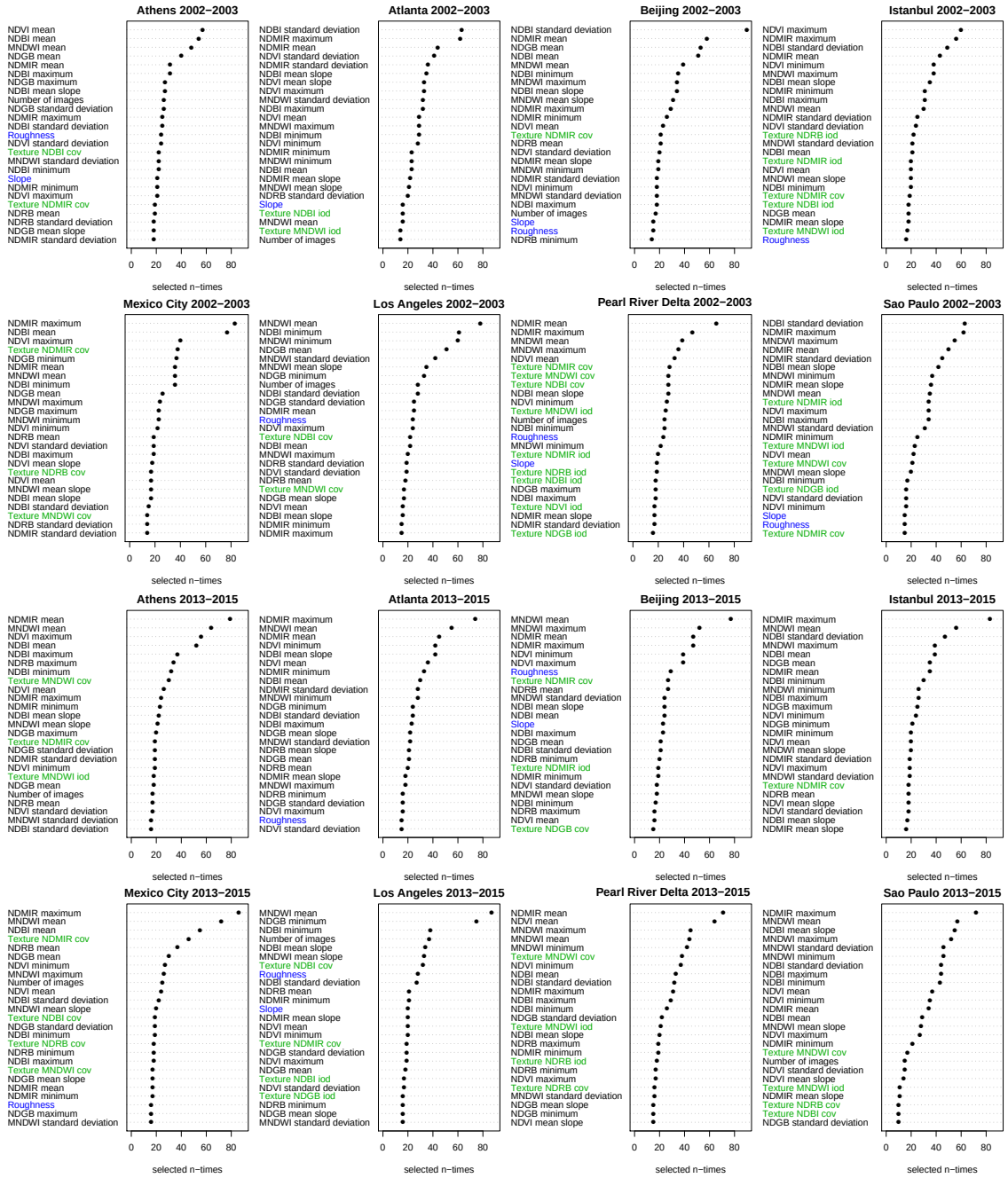

Figure 5.16: Outcome of feature selection that was applied on sampling configurations for each study site and investigation period.

but often also maximum, minimum and standard deviation have high occurrences. Nevertheless, one can notice that there is no clear agreement between all the graphics, which proves that in general all of them are relevant for the investigated classification purpose. In this context an exception is represented by the temporal mean MNDWI which ranks among the top 3 selected features in 11 over 16 experiments. Other indices that occur more frequently are the temporal maximum NDMIR and the temporal NDBI standard deviation. Despite they proved rather effective, texture features do not generally rank among the first places (however, this is also due to the fact that they help outlining low-density areas which occur less frequently in the analyzed test sites). Here, the coefficient of variation seems to be more effective than the index of dispersion. The texture features computed for MNDWI, NDBI and NDMIR are generally ranked higher.

6 Discussion

The obtained results proved that the proposed SVM based classification scheme is suitable to derive urban extent maps with very high accuracy. In the following sections an analysis is given on the data bases used, the classification results, and the possibility of transferring the developed method to other study areas.

6.1 Applicability of the Landsat based temporal statistics

The main objective of this study was a reliable delineation of the urban area extent in the periods 2002–2003 and 2013–2015 for the 8 cities included in the GEO SB-04 urban supersite initiative. For this purpose, temporal statistics extracted for different spectral indices derived from Landsat-5/7/8 data have been used (together with texture features as well as both slope and roughness calculated from the SRTM global 1 arc second DEM). Such an approach allows to characterize the behavior of different information classes over time and hence to better discriminate urban areas (which tend to be more stable) with respect to the other land cover classes experiencing a higher variability. However, so far only few studies investigated a similar strategy by means of multi-temporal images for monitoring urban areas (GRIFFITHS et al. 2010:434; SCHNEIDER 2012:702). Preliminary analysis carried out for assessing the capabilities of delineating urban areas based on manually labeled random samples derived for different information classes also shows the potential of employing these features for solving multi-class land cover classification problems. Just in few cases the presented approach tends to overestimate urban areas. This mostly occurs in correspondence of extraction and construction sites, which have under specific circumstances a comparable value range for the computed indices with respect to the urban areas.

Dealing with mass processing of multi-temporal data requires suitable resources and higher computational time. Nevertheless, in the new big-data era this is a relatively minor issue especially in the light of the current offer of cloud-computing facilities. The processing of Landsat data and the classification were performed using the GeoFarm processing infrastructure of DLR, hence, taking advantage of fast mass data processing.

As regards Landsat-5/7 data, the presence of jagged edges in the corresponding images negatively impacted the final temporal statistics as shown in the example reported in Figure 6.1. To solve this issue, after identifying the critical scenes by visual inspection, a suitable clipping has been applied prior to extracting the spectral indices.

It is worth pointing out that, theoretically, multi-temporal imagery collected over one-year period (hence covering a whole phenological cycle) might be sufficient for a proper discrimination of urban and non-urban areas. Nevertheless, due to the 16-day Landsat revisit time, this would statistically result in a low number of valid (i.e., cloud-free) scenes per pixel. Accordingly, the investigated time periods have been extended to 2 years, which

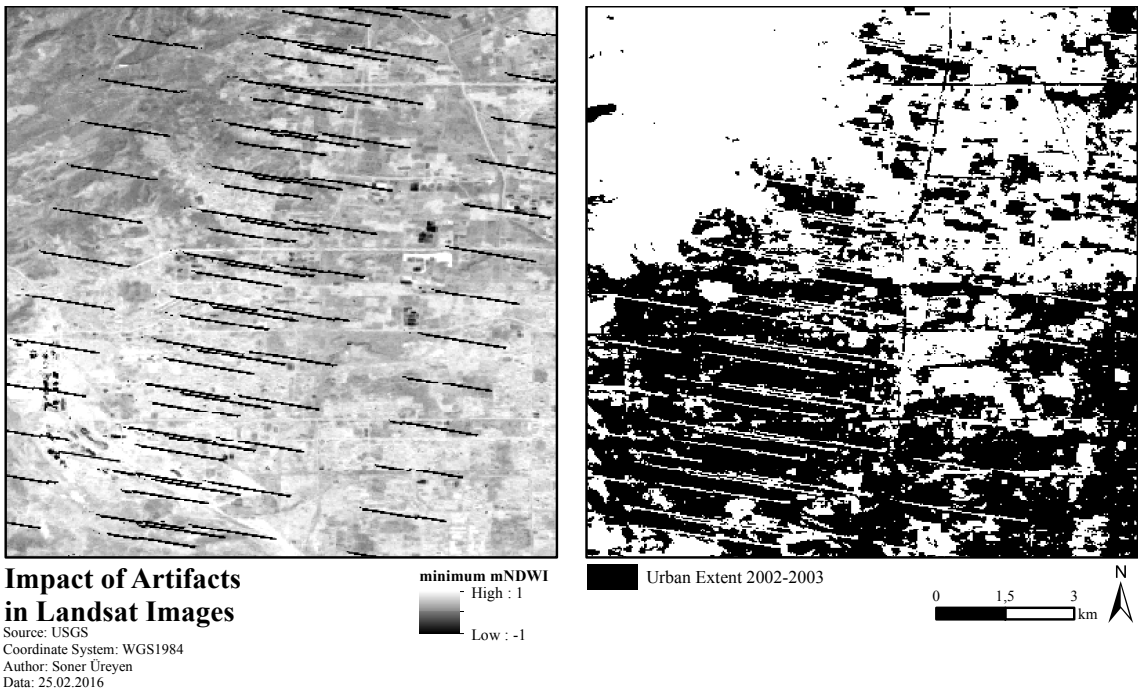

Figure 6.1: Exemplary depiction of the impact of the jagged edges on the temporal statistics and the resulting SVM based classification for the period 2002–2003. Here, the overlapping areas of the tiles 124032 and 123032 covering the study area Beijing are shown.

proved to be more robust and effective (only very new urban structures being built in the very specific time frame during which the Landsat images have been collected might not be included in the final classification maps).

Another aspect was the enhancement of the GUF by means of the temporal statistics to derive ‘training’ masks for an unsupervised collection of training points. The accuracy statistics of the ‘training’ masks as well as of the classification results clearly proved that the developed method is very effective and provides reliable data enabling an accurate automatic collection of training points. However, only few classification errors occurred mostly with regard to extraction sites probably due to noise of outliers in the training set.

6.2 Classification results

One of the objectives of this study was the implementation of a robust and accurate method suitable for outlining urban areas. To this aim first an unsupervised threshold based DT classifier has been used to assess the complexity of the considered problem and then a novel fully automatic SVM based system has been developed.

The qualitative and quantitative assessment of the classification maps derived by means of the unsupervised DT classifier proved in general the rather poor performances of such an approach for outlining urban areas with the extracted features. In particular, it exhibited either consistent under- or overestimation depending on the considered site.

In the literature, supervised classifiers proved significantly more effective. However, they require costly and time intensive collection of suitable training points, thus there is currently

high interest in automating such procedure (HUANG et al. 2015:16025). In this direction, a fully automatic approach based on SVM has been developed where training points for the urban and non-urban classes are derived from the GUF layer without the need for manual

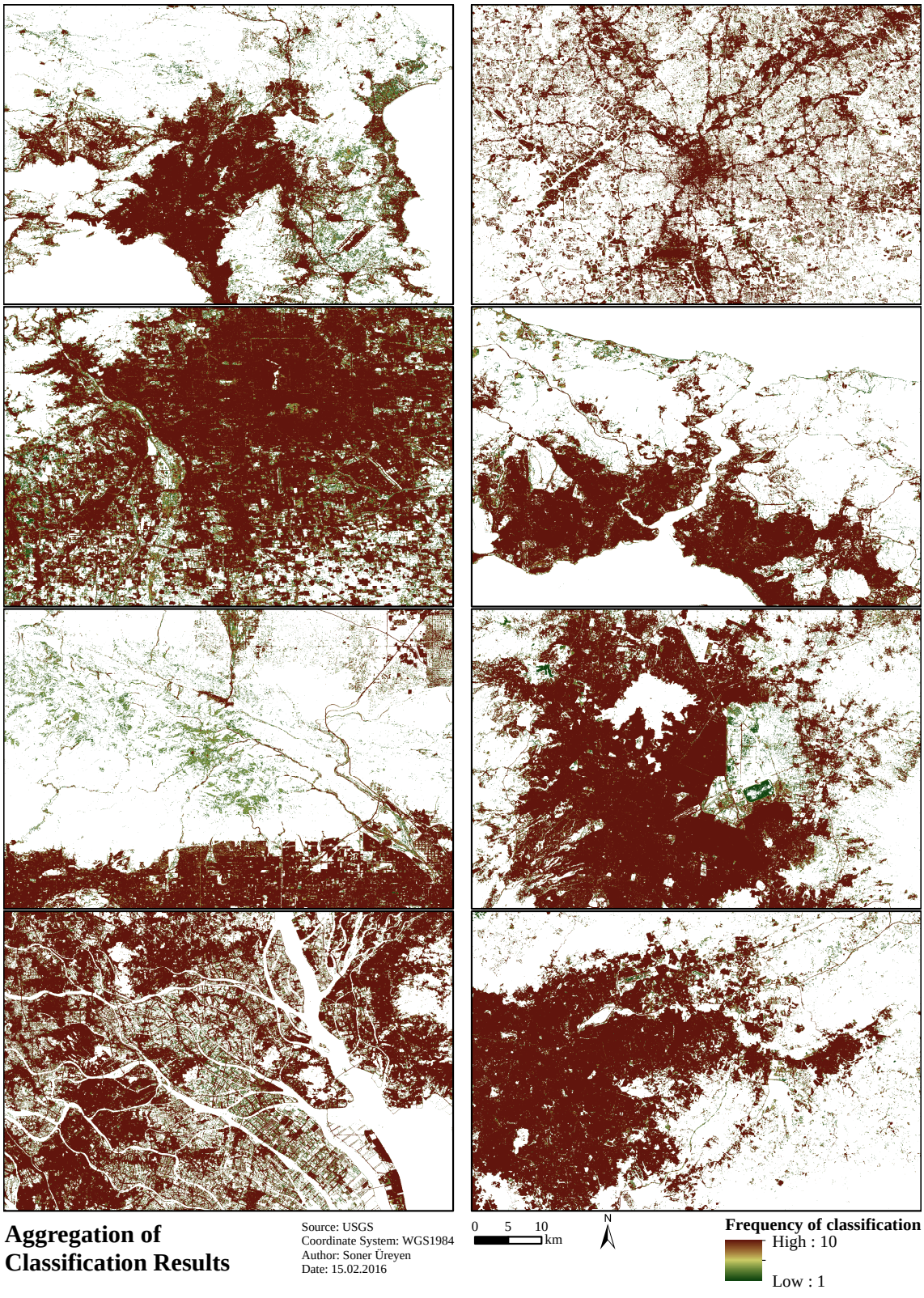

Figure 6.2: Subsets of study areas illustrating the aggregated classification results of configuration 4 that are used as input for the majority voting.

interaction. The obtained classification maps are very accurate for all the 8 analyzed AOIs, hence confirming the effectiveness of the presented method. Different configurations for selecting the training samples have been tested and compared and, finally, the one proving to be more robust is that where 1,000 labeled points are taken from both information classes. Performances proved stable against different random initializations of the training set; moreover, the employment of a majority voting approach allowed to further increase the classification accuracies. To this aim, in Figure 6.2 subsets of the chosen study areas are depicted where each pixel is associated with a different color depending on the number of times where it has been categorized as urban in the 10 classification maps obtained from as many random training set initialization with the optimal sampling configuration. Here it is interesting to notice how, in general, pixels appear either in darker red (associated with very high frequencies) or green (associated with very low frequencies) tones, while yellow ones (corresponding to intermediate frequencies) are basically missing. This once again proves how different training initializations tend to result in uncorrelated errors, which can then be overcome by using the majority voting strategy.

The derived accurate classification results for the 8 study areas allow to analyze the spatial and temporal dynamics of urban growth pattern within the investigated time frame by means of visual examination. A first inspection of the SVM based urban extent maps exhibits quite different patterns of urban growth among the study sites. Figure 5.15 reports that the Los Angeles site had the largest urban area extent (i.e., 5,134 km²) at the beginning of the investigated periods. Since 2002–2003, new developing urban areas contributed to an increase (i.e., 8.5 %) of approximately 400 km² until 2013–2015. The growth pattern for the Los Angeles area is primary characterized by new plots at the fringes of the urban extent of 2002–2003. The growth potentials in the Los Angeles area are highest for the southeastern land areas, since other possibilities are mostly limited by mountain ranges.

In contrast, the study areas in China exhibit by far the largest growth of urban areas. Here, the urban area extent of the Pearl River Delta study area grew by 114.1 % from 4,300 km² in 2002–2003, while it is possible to outline the extent of the single cities in 2002–2003 (Figure 5.13), namely Dongguan, Guangzhou, and Foshan, this is no longer the case in 2013–2015, where the cities grew to one megaregion. The growth pattern of this study site is characterized by extensive area increase along the fringes of the former urban extent; but, further remarkable increases in urban land cover are also pronounced by newly uprising urban patches in the south and northwest and also on artificial islands at the delta region. It is also possible to discern infilling (describes urban growth where a gap between existing urban areas is closed (Liu et al. 2010:672)) of urban areas in close distance to the urban centers of Guangzhou and Foshan. The vast number of construction sites indicate an ongoing trend of increase in urban area extent.

The urban area of the Beijing site reaches 5,421 km² in 2013–2015 from previously 2,680 km² (indicating a growth rate of 102.2 %). Figure 5.9 clearly shows a pattern of urban growth mostly occurring at the edges of the extent of 2002–2003. Beijing expanded

in the northwest along an existing trajectory to the city Changping. However, further growth in northern and western direction is bounded by mountain chains. Hence, as it is also depicted in Figure 5.9, the growth is pronounced in eastern and southern direction. In the southern part of the study area the growth is also characterized by fragmented growth, where new small urban areas emerged.

Comparatively, the increase of urban areas, in terms of absolute values, is not as high for the other study areas. Here, the highest growth rate (i.e., 71.5 %) was measured for Atlanta, where the urban area increased from 1,159 km² to 1,988 km². The growth is mainly marked by small sized developments along existing urban trajectories outside the administrative area, while within the borders of Atlanta city small infilling is visible (Figure 5.8).

Figure 5.12 reports urban growth in the Mexico City site occurred mainly in the northern and eastern part of the city. It is also noticeable that the city Toluca west of Mexico City exhibited considerable growth. The growth rate for this study area is 22.1 % (from 1,828 km² to 2,232 km²), while the newly developed urban areas are at the fringes of the former urban extent; but, there is also fragmented growth with small evolved plots in the north. The urban form of Mexico City and the areas where growth took place are caused by it's location in a valley surrounded by mountains.

Furthermore, the smallest increase in urban land cover is registered for the test site Sao Paulo. Indeed, the growth rate is only 1.8 % corresponding to an increase of 35 km² (i.e., 1,977 km² in 2013–2015). Here, growth occurred at very close distance to already existing urban areas.

The study areas with the smallest urban extent are Athens and Istanbul. In case of Istanbul, the urban extent is 1,140 km² in 2013–2015. It grew by 93 km² (8.9 %) with respect to the previous period. Figure 5.10 demonstrates most of the growth occurred on the European side of study area. The growth in Athens is primary characterized by new urban areas in far distance from the city center at fringes of sub-urban areas. The recent urban area extent for this test site is 767 km² (628 km² in 2002–2003).

6.3 Transferability of the classification methods

Another objective of this work was to develop a classification scheme which can be easily transferred to any new test site without the need for specific adjustments. The obtained classification accuracies prove that the DT classifier is not appropriate to this aim, while the SVM based system seems extremely suitable to the scope in the light of its effectiveness and robustness for the 8 analyzed study areas. In this context, the proposed novel approach has been further applied on four additional test sites, namely Addis Ababa (Ethiopia), Nairobi (Kenya), Kano (Nigeria) and Shanghai (China), for which almost all the currently existing global urban maps yielded rather low accuracies (WENG 2014a:5). Considering the transferability of the proposed SVM based approach a consistent size of 1x1° tiles has been selected for the introduced study areas. Additionally the results have been validated by generating a validation set of 400 points (equally split for the two classes) for each site.

Shanghai and Nairobi have been selected for investigating the capabilities of detecting very small size built-up areas in complex environments. In particular, the northern part of the Shanghai AOI is characterized by very small settlements developed along the borders of small agricultural fields, whereas in the Nairobi region the entire northern part of the study area is marked by mostly single housing patches arranged along roads with agricultural fields and plantations between them. The AOI north of the city Kano has been chosen to test the applicability of the proposed classification scheme on an area with only a limited number of small settlements. Additionally this AOI is also characterized by marginal or savanna vegetation. The study area Addis Ababa lies in a mountainous area and features a high share of agricultural and barren land as well as low residential areas which is the reason for it's inclusion.

From the analysis of the classification maps (Figure 6.3 and Figure 6.4) and the corresponding confusion matrices (Table 6.1) it is clear that the performances are in line with those exhibited for the 8 study areas from the GEO-SB04 supersite initiative. The highest accuracy measures are obtained for the study area Addis Abeba with an overall accuracy of 95.75 % and a Kappa of 0.915. In general, the overall accuracy and the Kappa coefficient are always higher than 91.50 % and 0.840. By comparing the derived classification results with corresponding VHR imagery, it becomes clear the proposed SVM based scheme is even capable of detecting urban areas in highly complex environments.

Table 6.1: Confusion matrices for the results of the additional study areas obtained from the SVM based classification for the period 2013–2015.

Addis Abeba		Reference		
		non-u.	urban	UA
Prediction	non-u.	198	2	99.00
	urban	15	185	92.50
	PA	92.96	98.93	

$OA = 95.75\ \%$; $\hat{K} = 0.915$

Nairobi		Reference		
		non-u.	urban	UA
Prediction	non-u.	193	7	96.50
	urban	27	173	86.50
	PA	87.73	96.11	

$OA = 91.50\ \%$; $\hat{K} = 0.840$

Kano		Reference		
		non-u.	urban	UA
Prediction	non-u.	194	6	97.00
	urban	17	183	91.50
	PA	91.94	96.83	

$OA = 94.25\ \%$; $\hat{K} = 0.885$

Shanghai		Reference		
		non-u.	urban	UA
Prediction	non-u.	195	5	97.50
	urban	20	180	87.50
	PA	90.70	97.30	

$OA = 93.75\ \%$; $\hat{K} = 0.875$

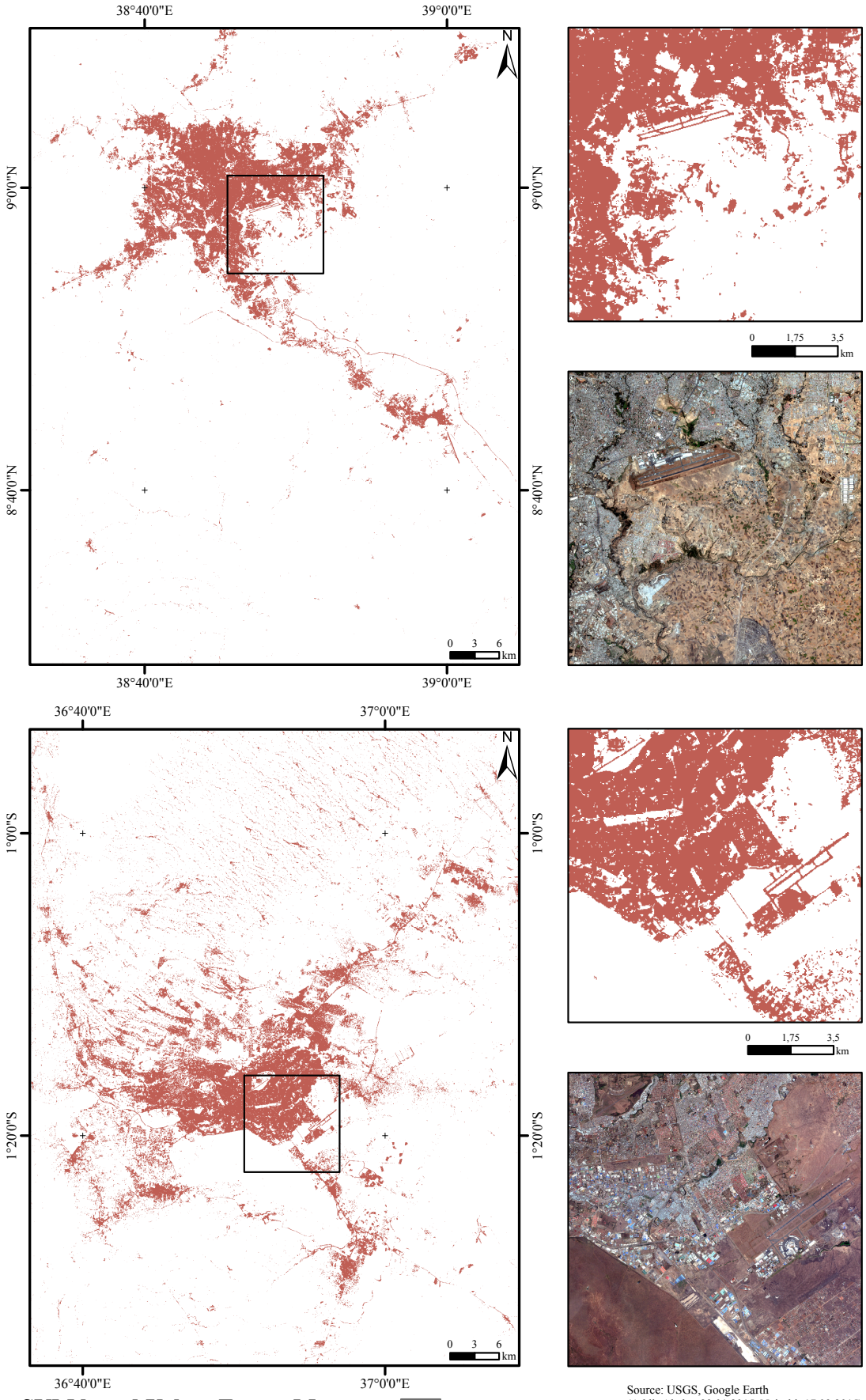

Figure 6.3: Urban extent maps computed for Addis Ababa and Nairobi to test the transferability of the proposed SVM based classification scheme.

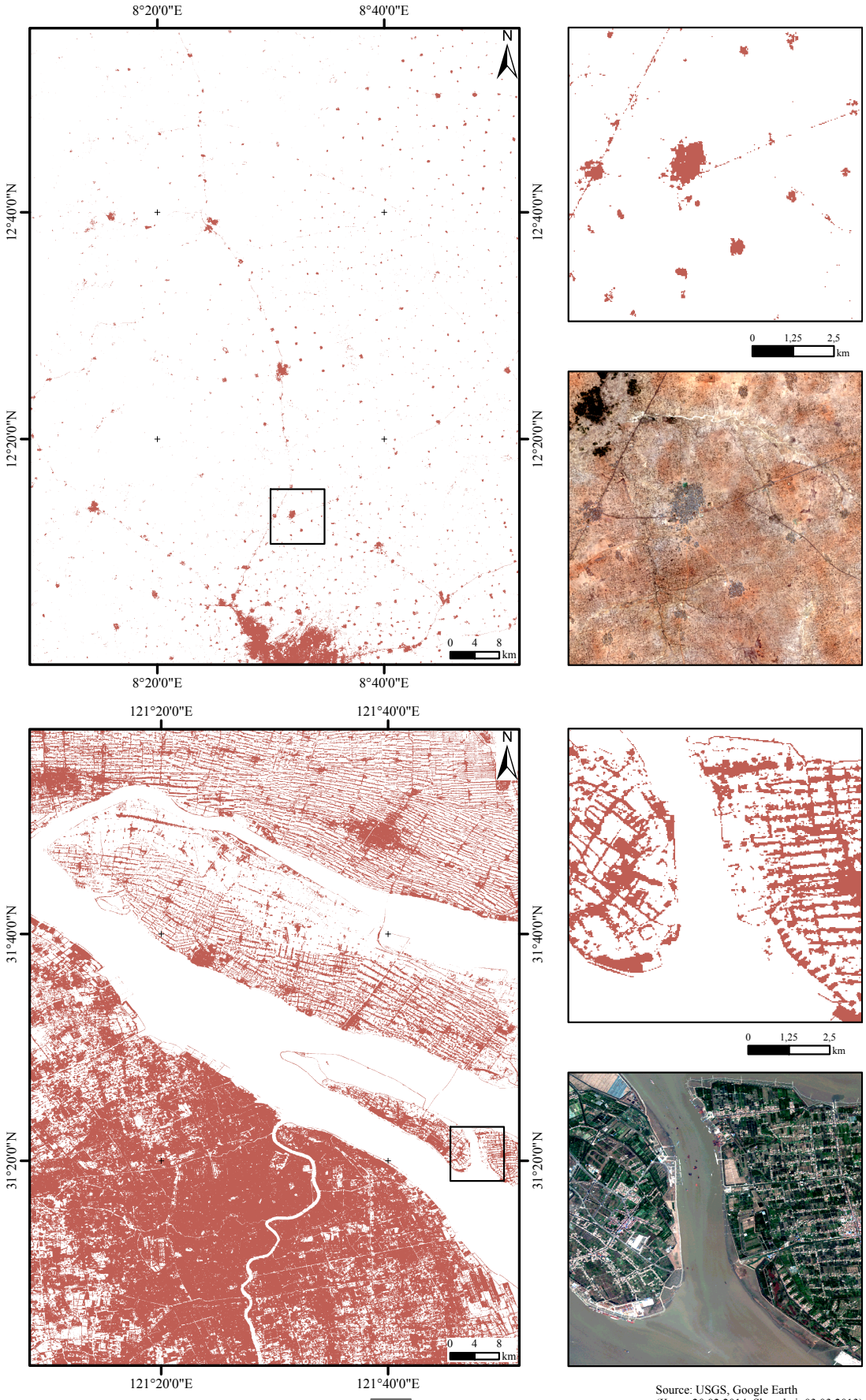

Figure 6.4: Urban extent maps computed for Kano and Shanghai to test the transferability of the proposed SVM based classification scheme.

7 Conclusions

In this thesis a novel classification system based on SVM has been proposed where the Landsat based temporal statistics are applied together with the GUF to delineate urban extent for the 8 sites of the urban supersites initiative and two investigation periods. The temporal statistics are derived for the time frame 2002–2003 and 2013–2015 and employed to enhance the GUF, which is implemented for automatic collection of training points. Since there is no rule for an optimal composition of the training set, 9 different sampling configurations of training data have been defined, where different numbers of points are tested for the urban and non-urban class. Furthermore, a total of 10 training sets have been collected per configuration to overcome possible impacts of the randomness resulting in as many urban extent maps. Afterwards, a majority voting strategy has been applied to each configuration to derive a final classification map selecting a pixel as urban if it is labeled as such in at least 6 over 10 instances. Next, an extensive validation has been carried out to assess the quality of the urban extent maps.

Here it has to be pointed out that the specified research objectives have been achieved with the following conclusions. One of the objectives was to enhance the GUF by means of the temporal statistics to enable an automated collection of a reliable training set. Indeed, the enhancement as well as some filtering approaches delivered highly accurate 'training' masks whereas, the overall accuracy and the Kappa coefficient for 2002–2003 are always better than 96.80 % and 0.936 and for 2013–2015 greater than 99.20 % and 0.968, respectively.

The main objective was to develop an automated classification scheme, which is capable of delivering consistent results and is robust as well as transferable on further study areas. Here, the classification procedure is automated implementing a shell script that solely needs the temporal statistics and the corresponding GUF and SRTM. The obtained overall accuracy and Kappa coefficient for both investigated periods are always higher than 91.38 % and 0.827, respectively, proving the consistency and robustness of the proposed classification system. Also the transferability of this approach has been tested on 4 additional study areas and here, too, the accuracy statistics are in line with the previously reported ones.

Furthermore, the high accuracy of the computed urban extent maps makes it possible to monitor and analyze urban growth patterns.

In the light of the Copernicus programme of the European Commission, a short prospect on future developments indicates large increase of satellite imagery particularly through the Sentinel-2 satellites. These will feature a considerably higher temporal resolution of 5 days and a slightly higher spatial resolution. The application of this data together with the Landsat data will allow the calculation of a more detailed and dense temporal statistics.

References

BAGAN, H. & YAMAGATA, Y. (2012): Landsat analysis of urban growth: How Tokyo became the world's largest megacity during the last 40 years. In: Remote Sensing of Environment 127, 210–222.

BANZHAF, E. & HÖFER, R. (2008): Monitoring Urban Structure Types as Spatial Indicators with CIR Aerial Photographs for a more Effective Urban Environmental Management. In: IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing 1, 2, 129–138.

BARTHOLOMÉ, E. & BELWARD, A. S. (2005): GLC2000: A new approach to global land cover mapping from Earth observation data. In: International Journal of Remote Sensing 26, 9, 1959–1977.

BRAZILIAN INSTITUTE OF GEOGRAPHY AND STATISTICS (2014): Cidades: Sao Paulo. http://cidades.ibge.gov.br/xtras/perfil.php?codmun=355030 (Accessed: 2015-12-12).

BURGES, C. J. (1998): A Tutorial on Support Vector Machines for Pattern Recognition. In: Data Mining and Knowledge Discovery 2, 2, 121–167.

CARTER, J. G., CAVAN, G., CONNELLY, A., GUY, S., HANDLEY, J. & KAZMIERCZAK, A. (2015): Climate change and the city: Building capacity for urban adaptation. In: Progress in Planning 95, 1–66.

CHEN, J., CHEN, J., LIAO, A., CAO, X., CHEN, L., CHEN, X., HE, C., HAN, G., PENG, S., LU, M., ZHANG, W., TONG, X. & MILLS, J. (2015): Global land cover mapping at 30m resolution: A POK-based operational approach. In: ISPRS Journal of Photogrammetry and Remote Sensing 103, 7–27.

CHEN, X.-L., ZHAO, H.-M., LI, P.-X. & YIN, Z.-Y. (2006): Remote sensing image-based analysis of the relationship between urban heat island and land use/cover changes. In: Remote Sensing of Environment 104, 2, 133–146.

CIESIN (2011): Global Rural-Urban Mapping Project, Version 1 (GRUMP v1): Urban Extents Grid. http://sedac.ciesin.columbia.edu/data/collection/grump-v1 (Accessed: 2016-01-14).

CIVCO, D. L. (1993): Artificial neural networks for land-cover classification and mapping. In: International Journal of Geographical Information Systems 7, 2, 173–186.

COHEN, J. (1960): A Coefficient of Agreement for Nominal Scales. In: Educational and Psychological Measurement 20, 1, 37–46.

CONGALTON, R. G. (1991): A review of assessing the accuracy of classifications of remotely sensed data. In: Remote Sensing of Environment 37, 1, 35–46.

CONGALTON, R. G. (2008): Sample Design Considerations. In: CONGALTON, R. G. & GREEN, K. [Eds.]: Assessing the Accuracy of Remotely Sensed Data: Principles and Practices 2nd edition 63–83. CRC Press, Taylor and Francis Group, Boca Raton.

CUI, L. & SHI, J. (2012): Urbanization and its environmental effects in Shanghai, China. In: Urban Climate 2, 1–15.

CULBERT, P. D., PIDGEON, A. M., ST.-LOUIS, V., BASH, D. & RADELOFF, V. C. (2009): The Impact of Phenological Variation on Texture Measures of Remotely Sensed Imagery. In: IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing 2, 4, 299–309.

DEVRIES, C., DANAHER, T., DENHAM, R., SCARTH, P. & PHINN, S. (2007): An operational radiometric calibration procedure for the Landsat sensors based on pseudo-invariant target sites. In: Remote Sensing of Environment 107, 3, 414–429.

DIJKSTRA, L. & POELMAN, H. (2014): A harmonised definition of cities and rural areas: the new degree of urbanisation. European Commission Brussels.

DUAN, Y., SHAO, X., SHI, Y., MIYAZAKI, H., IWAO, K. & SHIBASAKI, R. (2015): Unsupervised Global Urban Area Mapping via Automatic Labeling from ASTER and PALSAR Satellite Images. In: Remote Sensing 7, 2, 2171–2192.

ELVIDGE, C. D., TUTTLE, B. T., SUTTON, P. C., BAUGH, K. E., HOWARD, A. T., MILESI, C., BHADURI, B. & NEMANI, R. (2007): Global Distribution and Density of Constructed Impervious Surfaces. In: Sensors 7, 9, 1962–1979.

ESA (2016): GlobCover. http://due.esrin.esa.int/page_globcover.php (Accessed: 2016-01-15).

ESCH, T., MARCONCINI, M., FELBIER, A., ROTH, A., HELDENS, W., HUBER, M., SCHWINGER, M., TAUBENBOCK, H., MULLER, A. & DECH, S. (2013): Urban Footprint Processor - Fully Automated Processing Chain Generating Settlement Masks From Global Data of the TanDEM-X Mission. In: IEEE Geoscience and Remote Sensing Letters 10, 6, 1617–1621.

ESCH, T., TAUBENBÖCK, H., ROTH, A., HELDENS, W., FELBIER, A., THIEL, M., SCHMIDT, M., MÜLLER, A. & DECH, S. (2012): TanDEM-X mission — new perspectives for the inventory and monitoring of global settlement patterns. In: Journal of Applied Remote Sensing 6, 1, 1–21.

EUROSTAT (2015): Population on 1 January by age groups and sex - cities and greater cities. http://appsso.eurostat.ec.europa.eu/nui/show.do (Accessed: 2015-12-12).

FRIEDL, M. & BRODLEY, C. (1997): Decision tree classification of land cover from remotely sensed data. In: Remote Sensing of Environment 61, 3, 399–409.

GILL, S., HANDLEY, J., ENNOS, R. & PAULEIT, S. (2007): Adapting cities for climate change: the role of the green infrastructure. In: Built Environment (1978-) 30, 1, 97–115.

GRIFFITHS, P., HOSTERT, P., GRUEBNER, O. & VAN DER LINDEN, S. (2010): Mapping megacity growth with multi-sensor data. In: Remote Sensing of Environment 114, 2, 426–439.

GRIMMOND, S. (2007): Urbanization and global environmental change: local effects of urban warming. In: The Geographical Journal 173, 1, 83–88.

HALLEGATTE, S. (2009): Strategies to adapt to an uncertain climate change. In: Global Environmental Change 19, 2, 240–247.

HASTIE, T., TIBSHIRANI, R. & FRIEDMAN, J. (2009): The Elements of Statistical Learning. Springer Series in Statistics 2nd edition. Springer, New York.

HEIDEN, U., HELDENS, W., ROESSNER, S., SEGL, K., ESCH, T. & MUELLER, A. (2012): Urban structure type characterization using hyperspectral remote sensing and height information. In: Landscape and Urban Planning 105, 4, 361–375.

HEROLD, M., GARDNER, M. & ROBERTS, D. (2003a): Spectral resolution requirements for mapping urban areas. In: IEEE Transactions on Geoscience and Remote Sensing 41, 9, 1907–1919.

HEROLD, M., GOLDSTEIN, N. C. & CLARKE, K. C. (2003b): The spatiotemporal form of urban growth: measurement, analysis and modeling. In: Remote Sensing of Environment 86, 3, 286–302.

HUANG, B., XIE, C. & TAY, R. (2010): Support vector machines for urban growth modeling. In: GeoInformatica 14, 1, 83–99.

HUANG, C., DAVIS, L. S. & TOWNSHEND, J. R. G. (2002): An assessment of support vector machines for land cover classification. In: International Journal of Remote Sensing 23, 4, 725–749.

HUANG, X., WENG, C., LU, Q., FENG, T. & ZHANG, L. (2015): Automatic Labelling and Selection of Training Samples for High-Resolution Remote Sensing Image Classification over Urban Areas. In: Remote Sensing 7, 12, 16024–16044.

IMHOFF, M. L., ZHANG, P., WOLFE, R. E. & BOUNOUA, L. (2010): Remote sensing of the urban heat island effect across biomes in the continental USA. In: Remote Sensing of Environment 114, 3, 504–513.

KAUFMANN, R. K., SETO, K. C., SCHNEIDER, A., LIU, Z., ZHOU, L. & WANG, W. (2007): Climate Response to Rapid Urban Growth: Evidence of a Human-Induced Precipitation Deficit. In: Journal of Climate 20, 10, 2299–2306.

KITTLER, J., HATER, M. & DUIN, R. (1996): Combining classifiers. In: Proceedings of 13th International Conference on Pattern Recognition 20, 897–901.

KOTSIANTIS, S. B. (2013): Decision trees: a recent overview. In: Artificial Intelligence Review 39, 4, 261–283.

LALL, S. V., LOZANO-GRACIA, N., AGARWAL, O. P., DOWALL, D., KLEIN, M. & WANG, H. G. (2013): Planning, connecting, and financing cities - now : priorities for city leaders. World Bank Group, Washington, DC.

LI, X., GONG, P. & LIANG, L. (2015): A 30-year (1984–2013) record of annual urban dynamics of Beijing City derived from Landsat data. In: Remote Sensing of Environment 166, 78–90.

LI, X., ZHOU, W. & OUYANG, Z. (2013): Forty years of urban expansion in Beijing: What is the relative importance of physical, socioeconomic, and neighborhood factors? In: Applied Geography 38, 1–10.

LIU, X., LI, X., CHEN, Y., TAN, Z., LI, S. & AI, B. (2010): A new landscape index for quantifying urban expansion using multi-temporal remotely sensed data. In: Landscape Ecology 25, 5, 671–682.

LU, D., MAUSEL, P., BRONDIZIO, E. & MORAN, E. (2004): Relationships between forest stand parameters and Landsat TM spectral responses in the Brazilian Amazon Basin. In: Forest Ecology and Management 198, 1-3, 149–167.

MARCONCINI, M., METZ, A., ZEIDLER, J. & ESCH, T. (2015): Urban Monitoring in Support of Sustainable Cities. In: Proceedings of the 2015 Joint Urban Remote Sensing Event. Lousanne, Switzerland.

MASEK, J. G., LINDSAY, F. E. & GOWARD, S. N. (2000): Dynamics of urban growth in the Washington DC metropolitan area, 1973-1996, from Landsat observations. In: International Journal of Remote Sensing 21, 18, 3473–3486.

MASSON, V., MARCHADIER, C., ADOLPHE, L., AGUEJDAD, R., AVNER, P., BONHOMME, M., BRETAGNE, G., BRIOTTET, X., BUENO, B., DE MUNCK, C., DOUKARI, O., HALLEGATTE, S., HIDALGO, J., HOUET, T., LE BRAS, J., LEMONSU, A., LONG,

N., MOINE, M.-P., MOREL, T., NOLORGUES, L., PIGEON, G., SALAGNAC, J.-L., VIGUIÉ, V. & ZIBOUCHE, K. (2014): Adapting cities to climate change: A systemic modelling approach. In: Urban Climate 10, 407–429.

MCINERNEY, D. & KEMPENEERS, P. (2015): Open Source Geospatial Tools. Springer International Publishing, Cham.

MEGAHED, Y., CABRAL, P., SILVA, J. & CAETANO, M. (2015): Land Cover Mapping Analysis and Urban Growth Modelling Using Remote Sensing Techniques in Greater Cairo Region—Egypt. In: ISPRS International Journal of Geo-Information 4, 3, 1750–1769.

MERLÍN-URIBE, Y., CONTRERAS-HERNÁNDEZ, A., ASTIER-CALDERÓN, M., JENSEN, O. P., ZARAGOZA, R. & ZAMBRANO, L. (2013): Urban expansion into a protected natural area in Mexico City: alternative management scenarios. In: Journal of Environmental Planning and Management 56, 3, 398–411.

MILLER, R. B. & SMALL, C. (2003): Cities from space: potential applications of remote sensing in urban environmental research and policy. In: Environmental Science and Policy 6, 2, 129–137.

MOUNTRAKIS, G., IM, J. & OGOLE, C. (2011): Support vector machines in remote sensing: A review. In: ISPRS Journal of Photogrammetry and Remote Sensing 66, 3, 247–259.

NEMMOUR, H. & CHIBANI, Y. (2006): Multiple support vector machines for land cover change detection: An application for mapping urban extensions. In: ISPRS Journal of Photogrammetry and Remote Sensing 61, 2, 125–133.

PAL, M. & MATHER, P. M. (2003): An assessment of the effectiveness of decision tree methods for land cover classification. In: Remote Sensing of Environment 86, 4, 554–565.

PANDEY, B., JOSHI, P. & SETO, K. C. (2013): Monitoring urbanization dynamics in India using DMSP/OLS night time lights and SPOT-VGT data. In: International Journal of Applied Earth Observation and Geoinformation 23, 1, 49–61.

PAULEIT, S. & DUHME, F. (2000): Assessing the environmental performance of land cover types for urban planning. In: Landscape and Urban Planning 52, 1, 1–20.

PESARESI, M., EHRLICH, D., FERRI, S., FLORCZYK, A., FREIRE, S., HAAG, F., HALKIA, M., JULEA, A. M., KEMPER, T. & SOILLE, P. (2015): Global Human Settlement Analysis for Disaster Risk Reduction. In: ISPRS - International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences XL-7/W3, 837–843.

PESARESI, M., HUADONG, G., BLAES, X., EHRLICH, D., FERRI, S., GUEGUEN, L., HALKIA, M., KAUFFMANN, M., KEMPER, T., LU, L., MARIN-HERRERA, M. A., OUZOUNIS, G. K., SCAVAZZON, M., SOILLE, P., SYRRIS, V. & ZANCHETTA, L. (2013): A global human settlement layer from optical HR/VHR RS data: Concept and first results. In: IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing 6, 5, 2102–2131.

POTERE, D. & SCHNEIDER, A. (2009): Comparison of Global Urban Maps. In: GAMBA, P. & HEROLD, M. [Eds.]: Global Mapping of Human Settlement: Experiences, Datasets, and Prospects Chapter 13, 269–309. CRC Press, Taylor and Francis Group, Boca Raton.

PUDIL, P., NOVOVIČOVÁ, J. & KITTLER, J. (1994): Floating search methods in feature selection. In: Pattern Recognition Letters 15, 11, 1119–1125.

QUAN, B., BAI, Y., RÖMKENS, M., CHANG, K.-T., SONG, H., GUO, T. & LEI, S. (2015): Urban land expansion in Quanzhou City, China, 1995–2010. In: Habitat International 48, 131–139.

RIENOW, A. & GOETZKE, R. (2015): Supporting SLEUTH – Enhancing a cellular automaton with support vector machines for urban growth modeling. In: Computers, Environment and Urban Systems 49, 66–81.

RÖDER, A., PRÖPPER, M., STELLMES, M., SCHNEIBEL, A. & HILL, J. (2015): Assessing urban growth and rural land use transformations in a cross-border situation in Northern Namibia and Southern Angola. In: Land Use Policy 42, 340–354.

RONG YAN, YAN LIU, RONG JIN & HAUPTMANN, A. (2003): On predicting rare classes with SVM ensembles in scene classification. In: IEEE International Conference on Acoustics, Speech, and Signal Processing 3, 21–24.

ROUSE, J. W., HASS, R. H., SCHELL, J. & DEERING, D. (1973): Monitoring vegetation systems in the great plains with ERTS. In: Third Earth Resources Technology Satellite (ERTS) symposium 1, 309–317.

SCHNEIDER, A. (2012): Monitoring land cover change in urban and peri-urban areas using dense time stacks of Landsat satellite data and a data mining approach. In: Remote Sensing of Environment 124, 689–704.

SCHNEIDER, A., FRIEDL, M. A., MCIVER, D. K. & WOODCOCK, C. E. (2003): Mapping urban areas by fusing multiple sources of coarse resolution remotely sensed data. In: Photogrammetric Engineering and Remote Sensing 69, 12, 1377–1386.

SCHNEIDER, A., FRIEDL, M. A. & POTERE, D. (2009): A new map of global urban extent from MODIS satellite data. In: Environmental Research Letters 4, 4, 1–12.

SCHNEIDER, A., FRIEDL, M. A. & POTERE, D. (2010): Mapping global urban areas using MODIS 500-m data: New methods and datasets based on 'urban ecoregions'. In: Remote Sensing of Environment 114, 8, 1733–1746.

SCHNEIDER, A. & WOODCOCK, C. E. (2008): Compact, Dispersed, Fragmented, Extensive? A Comparison of Urban Growth in Twenty-five Global Cities using Remotely Sensed Data, Pattern Metrics and Census Information. In: Urban Studies 45, 3, 659–692.

SONG, C., WOODCOCK, C. E., SETO, K. C., LENNEY, M. P. & MACOMBER, S. A. (2001): Classification and Change Detection Using Landsat TM Data. In: Remote Sensing of Environment 75, 2, 230–244.

SVIREJEVA-HOPKINS, A., SCHELLNHUBER, H. J. & POMAZ, V. L. (2004): Urbanised territories as a specific component of the Global Carbon Cycle. In: Ecological Modelling 173, 2-3, 295–312.

TAM, B. Y., GOUGH, W. A. & MOHSIN, T. (2015): The impact of urbanization and the urban heat island effect on day to day temperature variation. In: Urban Climate 12, 1–10.

TAUBENBÖCK, H., ESCH, T., FELBIER, A., WIESNER, M., ROTH, A. & DECH, S. (2012): Monitoring urbanization in mega cities from space. In: Remote Sensing of Environment 117, 162–176.

THE WORLD BANK GROUP (2015): East Asia's Changing Urban Landscape: Measuring a Decade of Spatial Growth. The World Bank, Washington, DC.

TURKSTAT (2013): Population censuses. http://www.turkstat.gov.tr/PreTablo.do?alt_id=1047 (Accessed: 2015-12-10).

UNITED NATIONS (2015): World Urbanization Prospects: The 2014 Revision. Department of Economic and Social Affairs, Population Division New York.

U.S. CENSUS BUREAU (2012): Population Estimates: City and Town Intercensal Estimates (2000-2010). http://www.census.gov/popest/data/intercensal/cities/cities2010.html (Accessed: 2015-12-12).

U.S. GEOLOGICAL SURVEY (2014): SRTM 1 Arc-Second Global. Global Land Cover Facility University of Maryland.

U.S. GEOLOGICAL SURVEY (2015a): Frequently Asked Questions about the Landsat Missions. http://landsat.usgs.gov/best_spectral_bands_to_use.php (Accessed: 2016-02-01).

U.S. GEOLOGICAL SURVEY (2015b): Landsat Missions. http://landsat.usgs.gov/how_is_radiance_calculated.php (Accessed: 2015-01-20).

VILLA, P. (2012): Mapping urban growth using Soil and Vegetation Index and Landsat data: The Milan (Italy) city area case study. In: Landscape and Urban Planning 107, 3, 245–254.

WAI YEUNG YAN & SHAKER, A. (2014): Radiometric Correction and Normalization of Airborne LiDAR Intensity Data for Improving Land-Cover Classification. In: IEEE Transactions on Geoscience and Remote Sensing 52, 12, 7658–7673.

WAKODE, H. B., BAIER, K., JHA, R. & AZZAM, R. (2014): Analysis of urban growth using Landsat TM/ETM data and GIS—a case study of Hyderabad, India. In: Arabian Journal of Geosciences 7, 1, 109–121.

WALDE, I., HESE, S., BERGER, C. & SCHMULLIUS, C. (2014): From land cover-graphs to urban structure types. In: International Journal of Geographical Information Science 28, 3, 584–609.

WARMERDAM, F. (2008): The Geospatial Data Abstraction Library. In: Open Source Approaches in Spatial Data Handling Chapter 5, 87–104. Springer, Berlin.

WEEKS, J. R. (2010): Defining Urban Areas. In: RASHED, T. & JÜRGENS, C. [Eds.]: Remote Sensing of Urban and Suburban Areas 33–45. Springer Netherlands, Dordrecht.

WENG, Q. (2014a): What Is Special about Global Urban Remote Sensing? In: WENG, Q. [Ed.]: Global Urban Monitoring and Assessment through Earth Observation 1–12. CRC Press, Taylor and Francis Group, Boca Raton.

WENG, Q., ESCH, T., GAMBA, P., QUATTROCHI, D. & XIAN, G. (2014b): Global Urban Observation and Information: GEO's Effort to Address the Impacts of Human Settlements. In: WENG, Q. [Ed.]: Global Urban Monitoring and Assessment through Earth Observation 15–34. CRC Press, Taylor and Francis Group, Boca Raton.

WENG, Q., GAMBA, P., MOUNTRAKIS, G., PESARESI, M., LU, L., KEMPER, T., HEINZEL, J., XIAN, G., JIN, H., MIYAZAKI, H., XU, B., QURESH, S., KERAMITSOGLOU, I., BAN, Y., ESCH, T., ROTH, A. & ELVIDGE, C. (2014c): Urban Observing Sensors. In: WENG, Q. [Ed.]: Global Urban Monitoring and Assessment through Earth Observation 49–80. CRC Press, Taylor and Francis Group, Boca Raton.

WILSON, M. F. J., O'CONNELL, B., BROWN, C., GUINAN, J. C. & GREHAN, A. J. (2007): Multiscale Terrain Analysis of Multibeam Bathymetry Data for Habitat Mapping on the Continental Slope. In: Marine Geodesy 30, 1-2, 3–35.

WULDER, M. A., WHITE, J. C., GOWARD, S. N., MASEK, J. G., IRONS, J. R., HEROLD, M., COHEN, W. B., LOVELAND, T. R. & WOODCOCK, C. E. (2008): Landsat continuity: Issues and opportunities for land cover monitoring. In: Remote Sensing of Environment 112, 3, 955–969.

WULDER, M. A., WHITE, J. C., LOVELAND, T. R., WOODCOCK, C. E., BELWARD, A. S., COHEN, W. B., FOSNIGHT, E. A., SHAW, J., MASEK, J. G. & ROY, D. P. (2015): The global Landsat archive: Status, consolidation, and direction. In: Remote Sensing of Environment.

XIAO, P. F., WANG, X. H., FENG, X. Z., ZHANG, X. L. & YANG, Y. K. (2014): Detecting China's Urban Expansion Over the Past Three Decades Using Nighttime Light Data. In: IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing 7, 10, 4095–4106.

XU, H. (2006): Modification of normalised difference water index (NDWI) to enhance open water features in remotely sensed imagery. In: International Journal of Remote Sensing 27, 14, 3025–3033.

XU, H. (2007): Extraction of urban built-up land features from Landsat imagery using a thematic-oriented index combination technique. In: Photogrammetric Engineering and Remote Sensing 73, 12, 1381–1391.

XU, H., HUANG, S. & ZHANG, T. (2013): Built-up land mapping capabilities of the ASTER and Landsat ETM+ sensors in coastal areas of southeastern China. In: Advances in Space Research 52, 8, 1437–1449.

ZHA, Y., GAO, J. & NI, S. (2003): Use of normalized difference built-up index in automatically mapping urban areas from TM imagery. In: International Journal of Remote Sensing 24, 3, 583–594.

ZHOU, Y., YANG, G., WANG, S., WANG, L., WANG, F. & LIU, X. (2014): A new index for mapping built-up and bare land areas from Landsat-8 OLI data. In: Remote Sensing Letters 5, 10, 862–871.

ZHU, Z., WANG, S. & WOODCOCK, C. E. (2015): Improvement and expansion of the Fmask algorithm: cloud, cloud shadow, and snow detection for Landsats 4–7, 8, and Sentinel 2 images. In: Remote Sensing of Environment 159, 269–277.

ZHU, Z. & WOODCOCK, C. E. (2012): Object-based cloud and cloud shadow detection in Landsat imagery. In: Remote Sensing of Environment 118, 83–94.

A Discrimination analysis

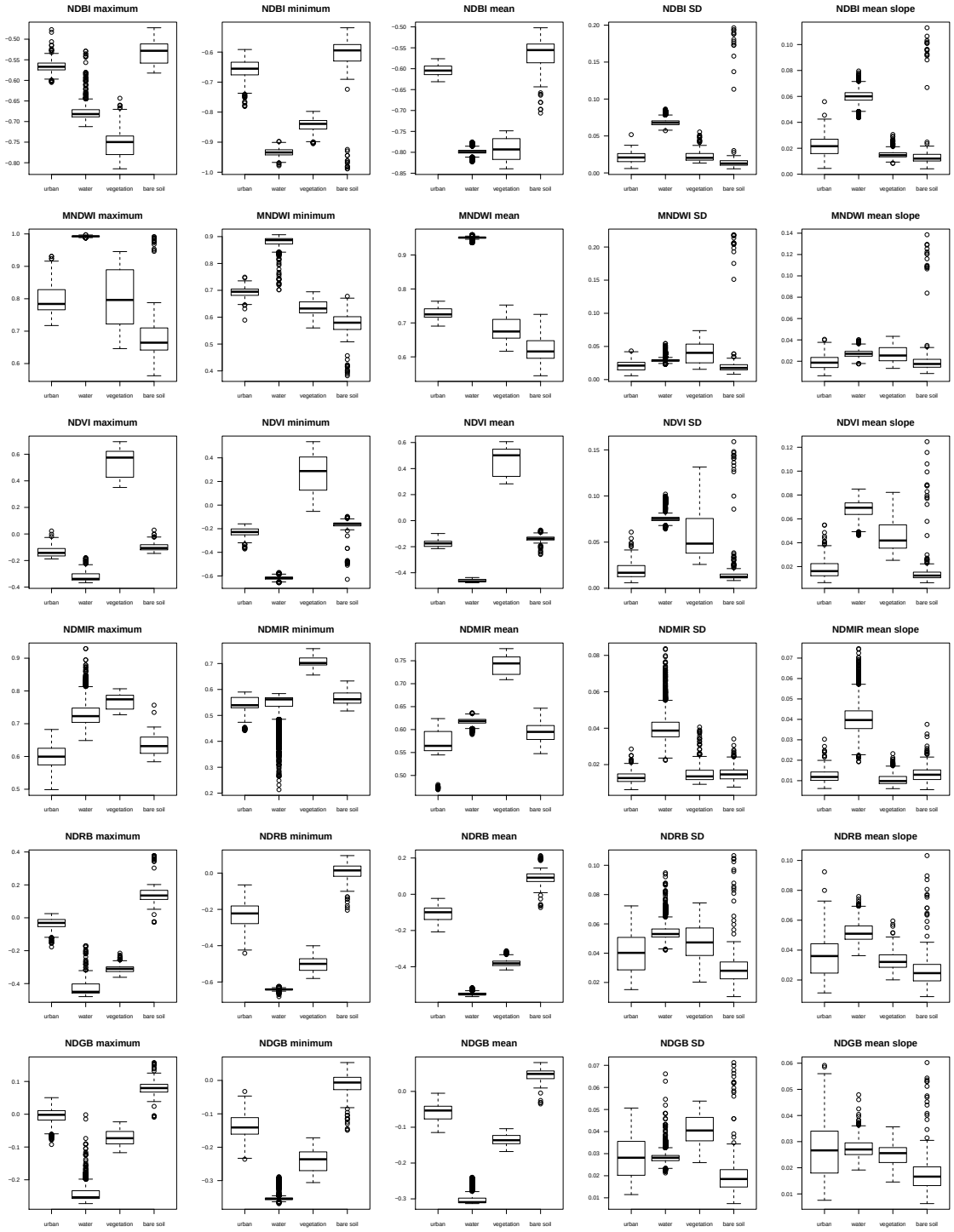

Figure A.1: Athens. A discrimination analysis was conducted for each study area comprising all spectral and temporal indices for the period 2013–2015 to investigate the effectiveness of the temporal statistics.

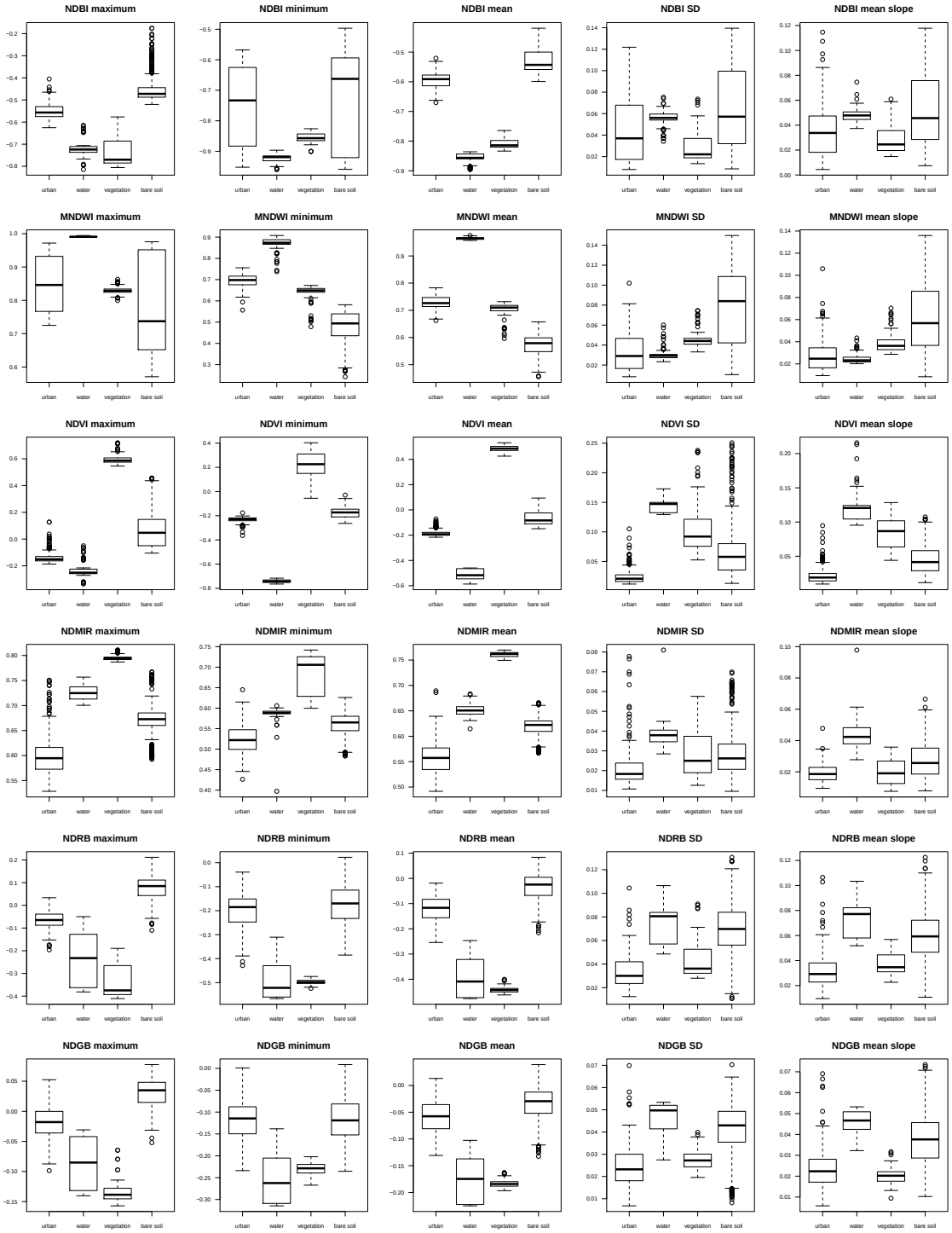

Figure A.1 (continued)**:** Atlanta.

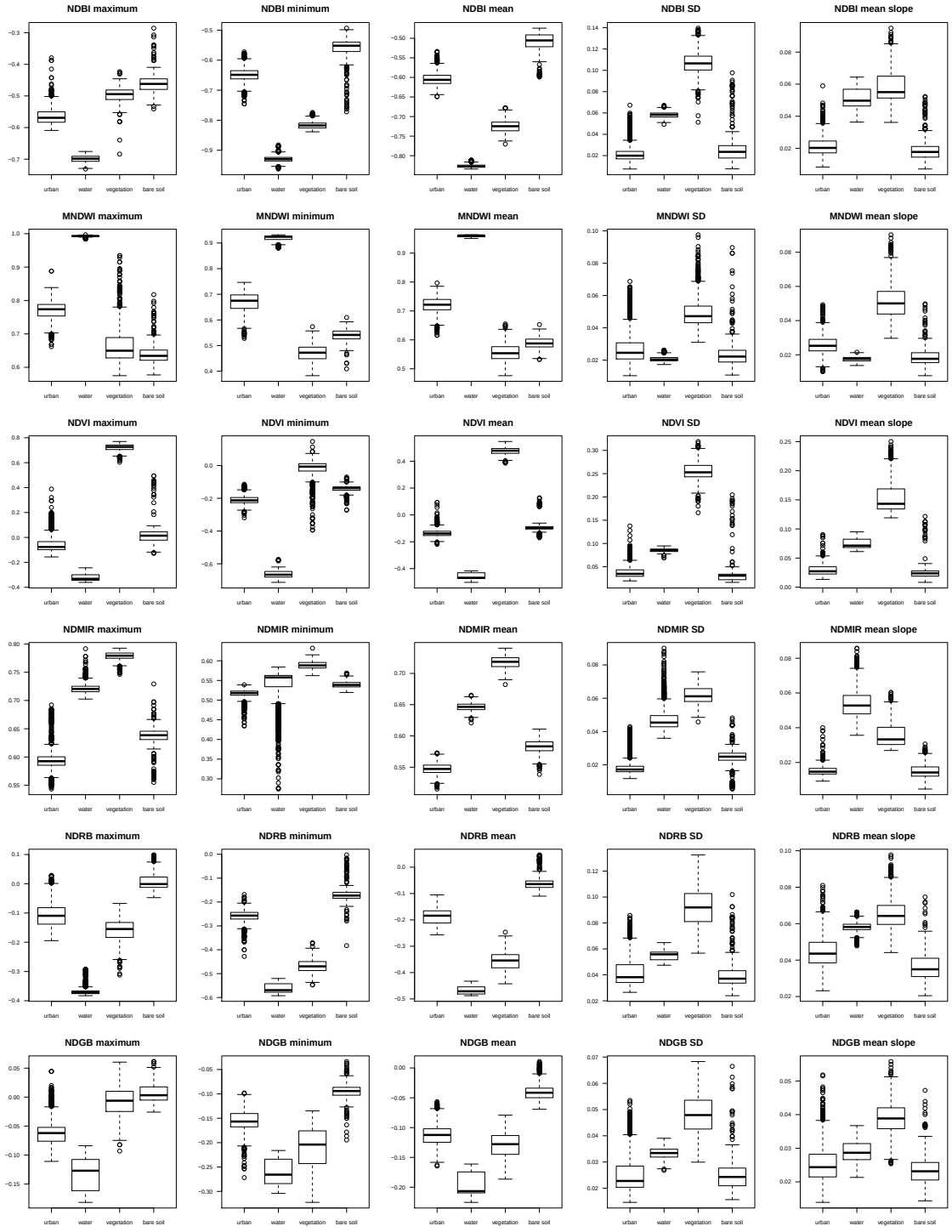

Figure A.1 (continued)**:** Beijing.

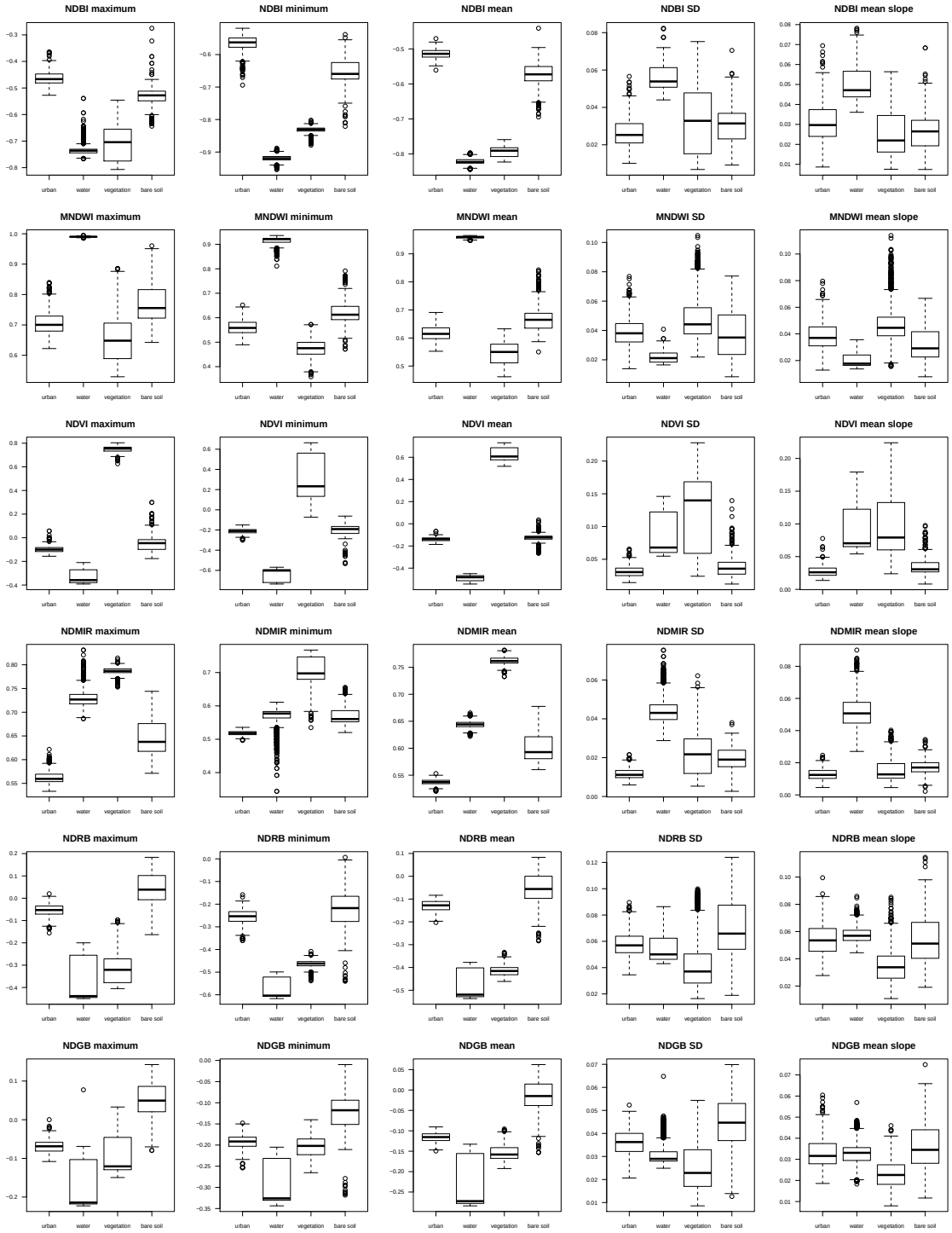

Figure A.1 (continued)**:** Istanbul.

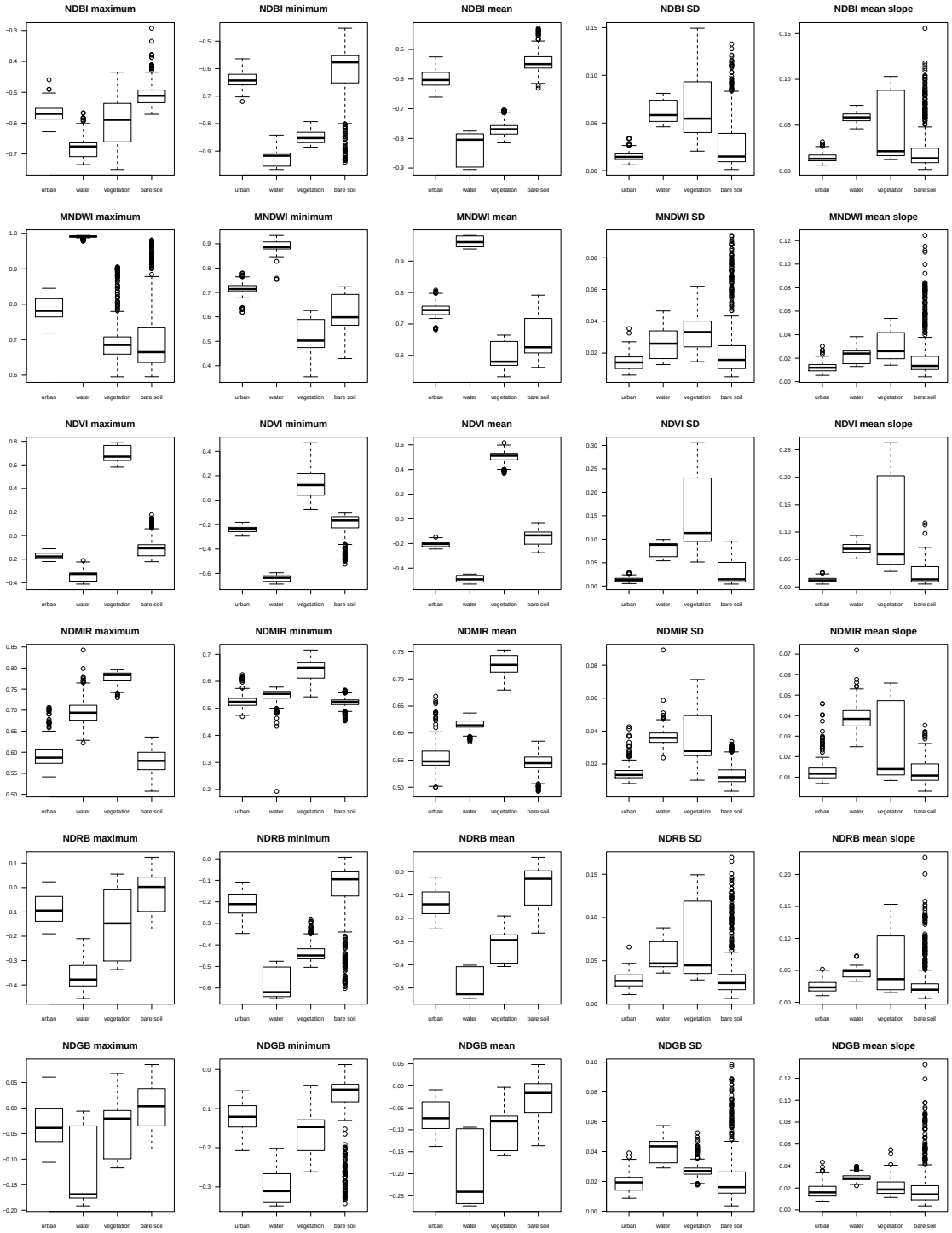

Figure A.1 (continued)**:** Los Angeles.

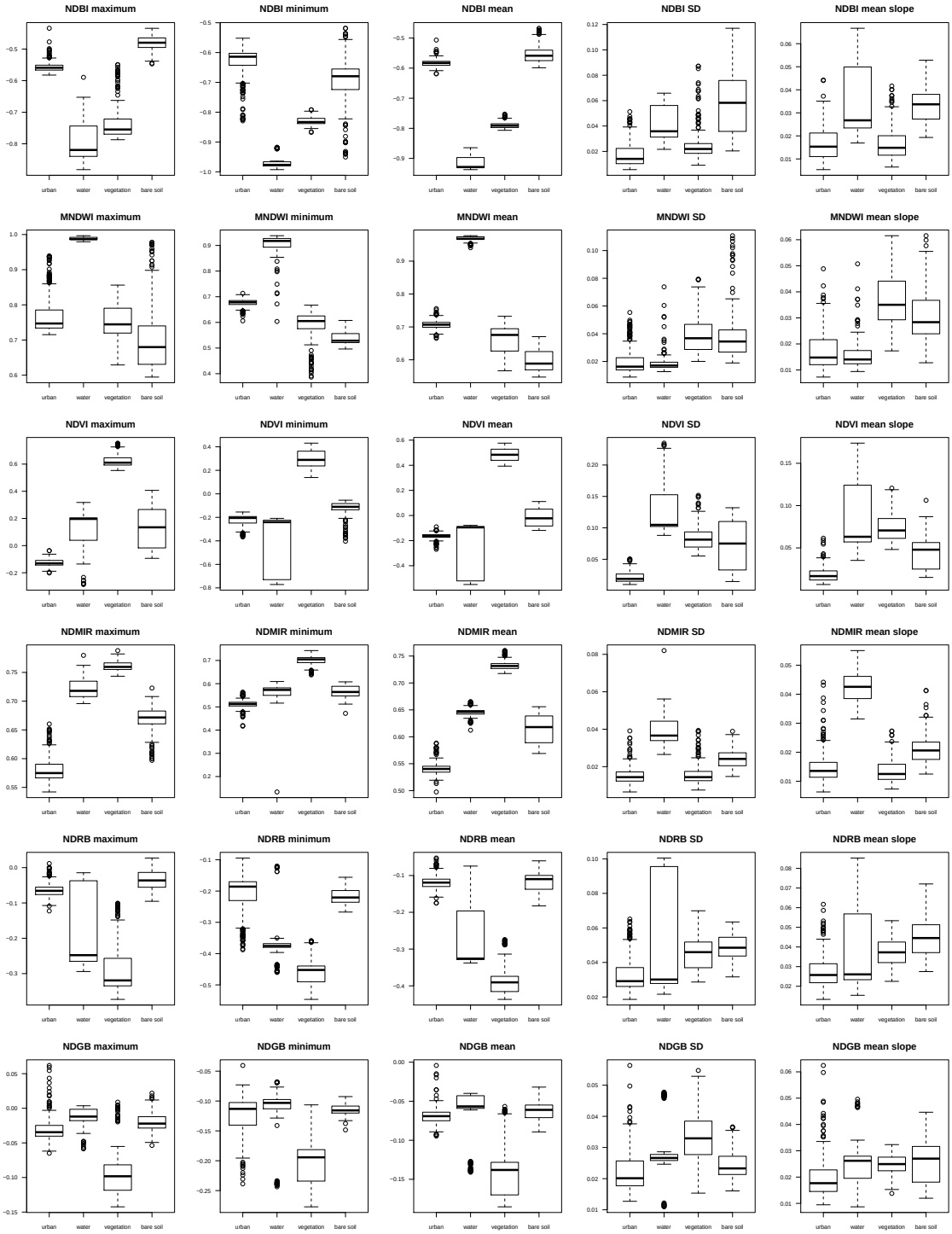

Figure A.1 (continued)**:** Mexico City.

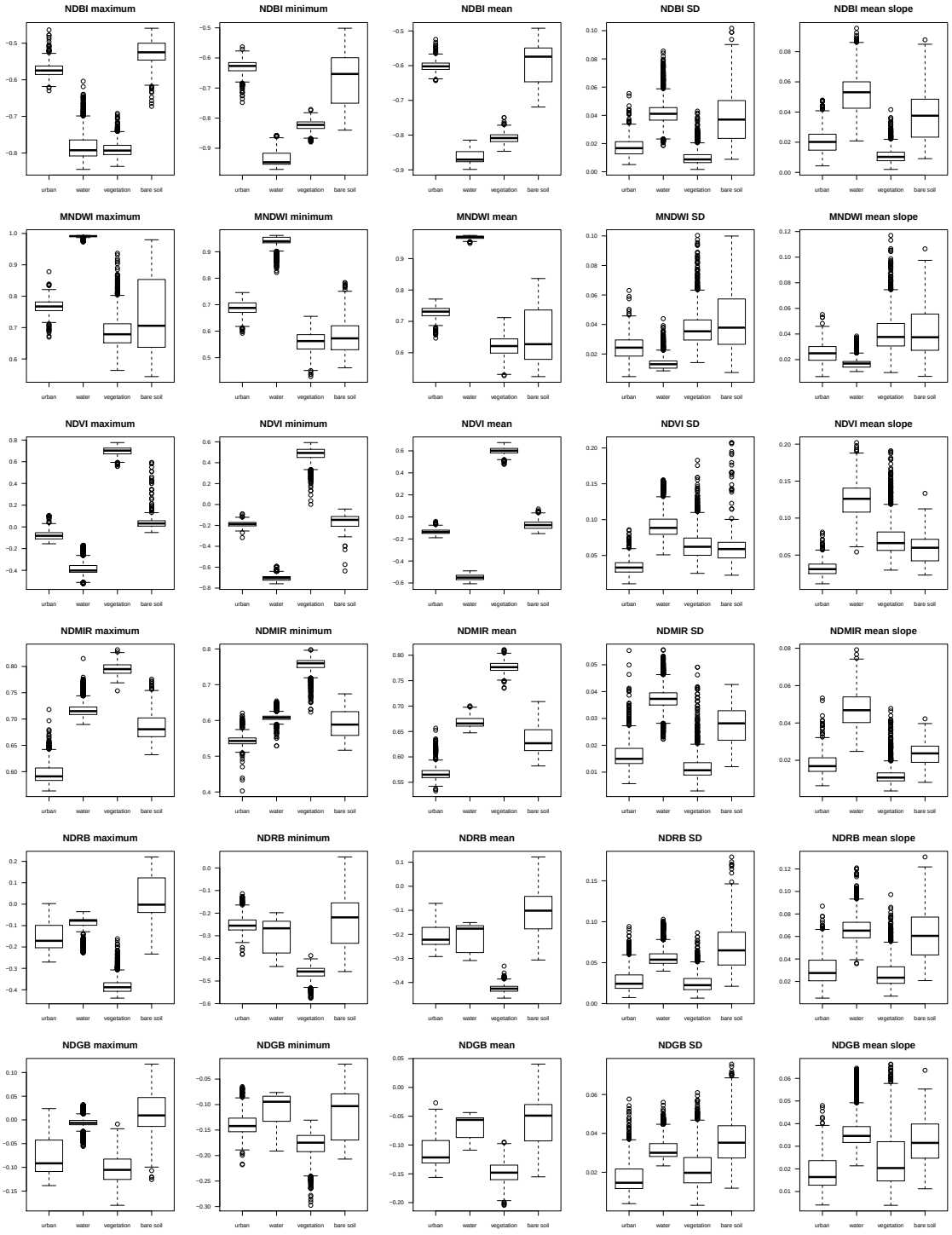

Figure A.1 (continued)**:** The Pearl River Delta.

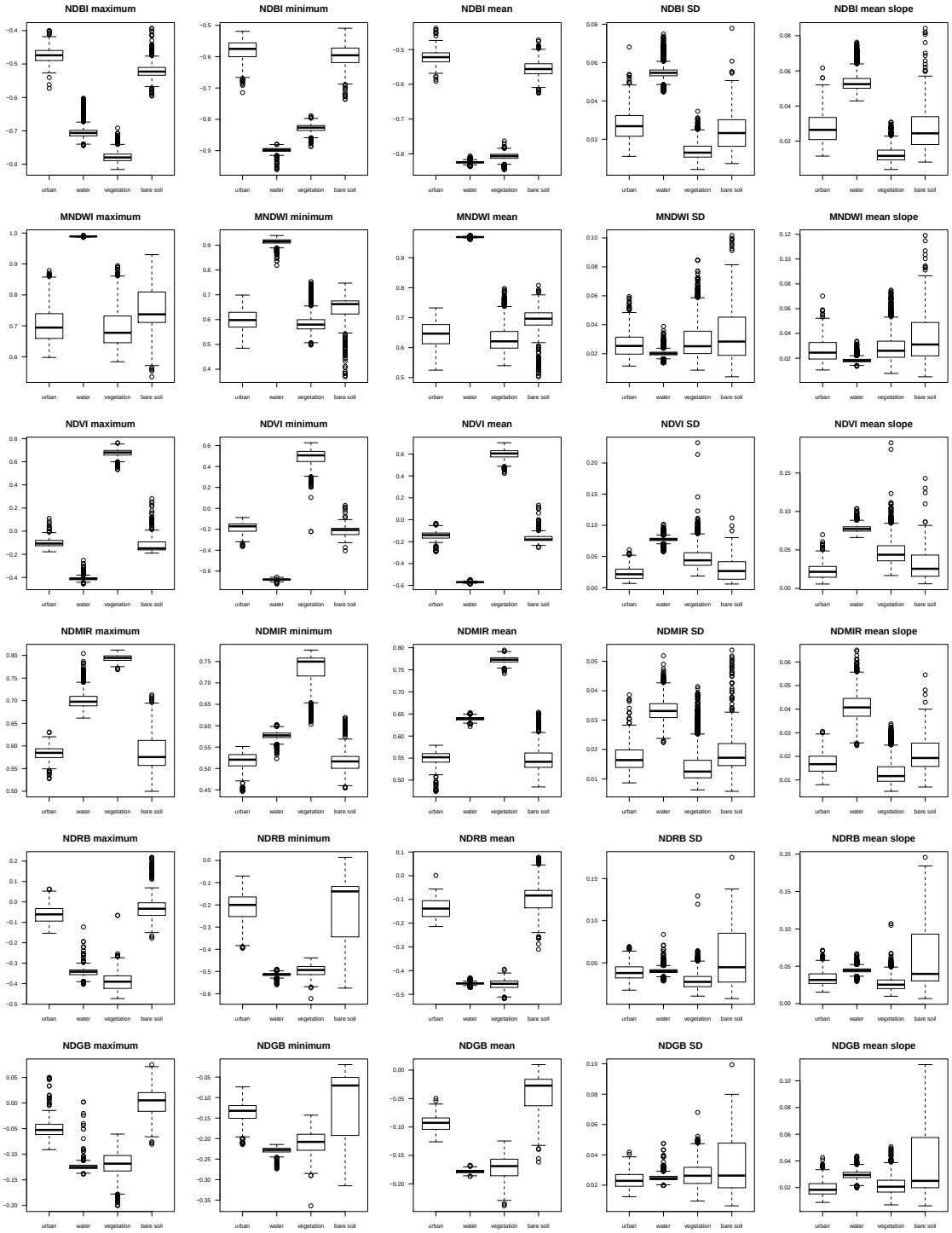

Figure A.1 (continued)**:** Sao Paulo.